مقـــدمة فـــى

الإحصاء الاجتماعى

تأليف

أ. د. اعتماد محمد علام
أستاذ علم الاجتماع
ووكيل كلية البنات للدراسات العليا
والبحوث بجامعة عين شمس

الناشـــر
مكتبة الأنجلو المصرية
١٦٥ ش محمد فريد - القاهرة

فهرس الكتاب

إهـداء

إلي شريكي في رحلة العمر . . .

من شد ازري في طلبي للعلم . الي زوجي مع

باقـة من الياسـمين المـندي بقطـرات النـدي

اعتماد علام



* مقدمـــة *

يهدف هذا المؤلف إلى محاولة عرض الأساليب الأساسية للإحصاء الوصفى بسهولة ويسر ، دون إقحام الطالب المبتدىء فى عمليات رياضية مركبة .

فيما يختص بأهمية الإحصاء ، يكفى القول أنها أصبحت جزءاً من لغة الخطاب اليومى على مستوى الجمهور ، بالإضافة إلى فوائدها الجمة لمختلف مجالات العلوم سواء على مستوى النظريات أو البحوث الميدانية .

على مستوى المجتمع ، يتحدث الأفراد بلغة الأرقام عند سؤالهم عن درجة حرارة الجو أو عن انخفاض وارتفاع مؤشر بورصة الأوراق المالية . كما يطالعون فى الصحف اليومية الأشكال البيانية والجداول الخاصة بالميزانيات التى تعرض بياناتها فى عرض موجز ، يوضح ماتحمله من مكاسب أو خسائر . بالمثل يشاهد الجمهور على شاشات التليفزيون بيانات إحصائية تتعلق بمعدل المواليد وأخرى تتعلق بالمد والجزر ، والنسب المئوية لحوادث السيارات على الطرق وداخل المدن ..... الخ .

على المستوى الأكاديمى ، لم تعد الإحصاء جزءاً أساسياً من علم الرياضيات فقط ، بل تدخل الإحصاء فى جميع فروع العلم . فمن خلال النماذج الإحصائية ، تحاول العلوم الانسانية تحويل المقولات الى فروض ، ثم مؤشرات يمكن قياسها إمبيريقيا ، كما تفيد الإحصاء فى تعلم الباحث الأسلوب الصحيح لقراءة الجداول وتلخيص البيانات وتفسير نتائج البحث الامبيريقى .

من المنظور الوظيفى ، تنقسم الإحصاء الى قسمين أساسيين أولهما يعرف بالإحصاء الوصفى الذى يعتبر مجال إهتمامنا الأساسى فى هذا المؤلف ، وثانيهما الاحصاء الاستدلالى ، والذى سوف يكون محور إهتمام مولفنا الثانى بإذن الله .

تتعدد أساليب الإحصاء الوصفى فى تعاملها مع البيانات حول ظاهرة ما محل الدراسة ، ووفقاً للأهداف التى يسعى الباحث فى تحقيقها . ونحاول فى هذا

المؤلف شرح الخطوات الأساسية المتبعة فى استخدام كل أسلوب من أساليب الإحصاء الوصفى وذلك فى عرض موجز وبسيط ، مع تدعيم الشرح بأمثلة محلولة وتوضيح مزايا وعيوب كل أسلوب ، حتى يستطيع الباحث أن يختار الأسلوب المناسب لنوعية البيانات التى تم جمعها . ولتحقيق هذا الهدف، رأينا أن نقسم هذا المؤلف الى ستة فصول بحيث يضم كل فصل عدداً من أساليب الإحصاء الوصفى .

يضم الفصل الأول التعريف بالإحصاء الاجتماعى وأهميته ، ويختص الفصل الثانى بمستويات القياس والعرض الجدولى للبيانات . ويتناول الفصل الثالث التمثيل البيانى للبيانات ، ويختص الفصل الرابع بمقاييس النزعة المركزية ، ويحتوى الفصل الخامس على مقاييس التشتت وتختتم هذا المؤلف بالفصل السادس الذى يضم الارتباط والانحدار الخطى .

وأرجو أن يجد القارىء المتعة فى دراسته لهذا المؤلف بالقدر الذى سعدت به أثناء كتابته .

والله ولى التوفيق ،

المؤلفة

إعتماد محمد علام

القاهرة اكتوبر ١٩٩٨ .

الفصل الأول

الإحصاء الاجتماعي : التعريف والأهمية

مقدمـــــة :

١ – تعريف الاحصاء .

٢ – الأساليب الإحصائية .

٣ – تعريف البيانات ومصادرها .

٤ – المجتمع الأصلى والعينة .

٥ – تصنيف المتغيرات .

٦ – مراحل الاختيار الاحصائى فى البحوث الاجتماعية .

٧ – التعرف على استخدام الآلة الحاسبة فى مجال الإحصاء .

٨ – وظائف الإحصاء فى البحوث الاجتماعية .



الفصل الأول

الإحصاء الاجتماعى :
التعريف والأهمية

مقــدمة :

إذا تأمل الإنسان كل مايدور حوله من أحداث وتغيرات ومعلومات مقروءة أو مرئية أو مسموعة حتى ما يتضمن خطاب الحياه اليومية ، سوف يجد نفسة محاطاً بالبيانات الإحصائية . فيوميا تطالعنا الصحف ببيانات إحصائية فى شكل جداول أو رسومات بيانية أو نسب مئوية حول البطالة فى سوق العمل، أو تصاعد أسهم فى بورصة الأوراق المالية، أو نسبة الحوادث ومعدلاتها علي الطرق خلال عام أو خلال فترة زمنيةمعينة ...... الخ. حتى فيما يدور من حديث فى خطاب الحياة اليومية نتحدث عن انخفاض الأسعار حول سلعة معينة أو النسب المئوية لمجموع درجات الطلاب والطالبات فى امتحان الثانوية العامة ...... إلخ. من ثم نقول إن الاحصاءات أصبحت جزءا هاماً من حياة الانسان ، إلاأنها قد تشير إلى موضوعات مختلفة من خلال رؤية وأساليب متباينة بين فرد وأخر وبين باحث فى مجال علمى ما أو باحث فى مجال علمى آخر. فمثلا، يناقش خبراء الطقس الإحصائيات اليومية حول ارتفاع درجة حرارة الجو أو انخفاضها واحتمالات سقوط الأمطار ونسبة كثافتها خلال الأيام القادمة ، وسرعة الرياح ، والنسب المئوية لرطوبة الجو ، وحالة البحر من مد وجذركل ذلك فى شكل إحصائيات وصفية لما تم رصده بالفعل عن طريق أجهزة الرصد والقياس ، واستخدام الإحتمالات فى توقع الاحوال الجوية المستقبلية حتى باستخدام الأقمار الصناعية التى تعتمد اعتماداً أساسيا على البيانات الإحصائية، وينطبق هذا القول علي خبراء الرياضة حيث يستخدمون النسب والبيانات الإحصائية فى الوصف والتعليق علي مباريات كرة القدم . من جهة أخرى، يختلف أسلوب الخطاب الإحصائى للباحثين فى العلوم الإنسانية والفيزيقية عنة للفئات التى أشرنا اليها فى المثال السابق. فالباحث من خلال ما يجمعة من بيانات حول ظاهرة معينة أو متغيرما، يبحث

عن الأدوات الإحصائية الملائمة لتحليل هذه البيانات. أما المشتغلون بالعلوم الرياضية فإنهم يصفون الإحصاء كجزء أساسى من علوم الرياضيات .

من المنظور التطبيقى، نجد أن مجال الإحصاء يعم مختلف التخصصات العلمية رغم التباين فيما بينها من طب، صحة عامة، وادارةالاعمال، علم الاجتماع وعلم النفس ....إلخ.

من المنظور الوظيفى، ينقسم الإحصاء الاجتماعى إلى قسمين اساسين أولهما الاحصاء الوصفى Discriptive Statistics وهذا مانهتم به فى هذا المؤلَف . ثم الاحصاء الاستدلالى inferential statistics التى تصف المجتمع الأصلى باستخدام معلومات من عينات صغيرة نسبيا تمثل هذا المجتمع . وتعتمد معظم البحوث الاجتماعية على الإحصاء الاستدلالى لأن دراسة المجتمعات الأصلية كبيرة الحجم عادة ما تكون صعبة وتكلفة دراستها عالية (Kurtz,1983:2) .

نظراً للتوسع فى استخدامات الحاسب الآلى فى البحث الاجتماعى، كان ضرورياً على الباحثين تعليم لغة خاصة يستخدمونها فى الإتصال بالحاسب الآلى ومن أكثر اللغات الاجتماعية استخداماً وانتشاراً لتحقيق هذا الإتصال ما يعرف بالحزم الإحصائية للعلوم الاجتماعية Statistical Package for the Social Sciences وتعرف اختصارا بلغة (SPSS) ويمكن للطلاب والباحثين اسخدام هذه اللغة مع توخى الحذر الذي يتطلب منهم التفهم الجيد للتعريفات والمتغيرات الخاصة بهذه اللغة مع القدرة على انتهاج الأسلوب المناسب لتحليل البيانات ومعنى الدلالات الإحصائية للنتائج بعد تحليلها بالحاسب الآلى(Klecka et al, 1975:1 & Nei et al, 1975: 3) .

تستخدم الإحصاء مستويات للقياس وتتعامل فى العلاقات بين المتغيرات على اختلاف أنواعها. كما تتدخل الاحصاء فى الربط بين المفهوم النظرى والتعريف الإمبريقى لمتغيرماً من خلال ما يعرف بالنماذج Models. ونجد فى هذا الصدد من يرى إن لكل متغير مفهوم وتعريف معا. فمن خلال استخدام أساليب بحثية معينة وصياغة الفروض واتخاذ قرارات من جانب الباحثين يمكن من خلال استخدام النماذج الإحصائية تحويل المقولات النظرية لمتغير ما إلى أبعاد

Dimensions يمكن قياسها إمبريقيا وتصميم نموذج إحصائي مع إمكانية تعريف هذا المتغير إمبريقيا.

تأسيسا على ماسبق وانطلاقا منة تنقسم المناقشة فى هذا الفصل إلى:

١ – تعريف الإحصاء.

٢ – الأساليب الإحصائية (اقسام الإحصاء).

٣ – تعريف البيانات ومصادرها.

٤ – المجتمع الاصلى والعينة.

٥ – تصنيف المتغيرات.

٦ – مراحل الاختبار الإحصائى فى البحوث الاجتماعية.

٧ – التعرف على استخدام الآلة الحاسبة فى مجال الإحصاء.

٨ –وظائف الإحصاء فى البحوث الاجتماعية.

تعريف الإحصاء

فيما يتصل بتعريف الإحصاء، قد لانجد إتفاقاً تاماً وصريحاً بين المشتغلين بها حول تعريف كامل ومتفق عليه الا أن هناك أسساً جوهرية ليست موضع اختلاف بينهم وتمثل فى مضمونها تعريفا للاحصاء (Blalock , 1972: 4).

تعرف الإحصاء بمجموعة الطرق التى تستخدم فى تجميع ووصف وتحليل البيانات الرقمية الدالة على جوانب متعددة ومتباينة للحياة الاجتماعية كما تعرف الإحصاء بأنها تشكيله من النظرية والمناهج يتم تطبيقها بغرض فهم البيانات وتقديم دلالة امبريقية على قبول أورفض النظريات المستخدمة فى العلوم السلوكية.

الأســاليب الاحصــائية:

تنقسم الطرق أو الأساليب الإحصائية الى قسمين أساسيين هما:

١ – **الإحصاء الوصفى** ويتالف من مجموعة الأساليب التى تصف الظواهر الاجتماعية من خلال أوصاف رقمية. فمثلا إذا أردت أن تصف مجتمعك المحلى الذى تعيش بداخلة بدلالة ثلاثة متغيرات هى النوع Sex، العمر Age،

ودخل الأسرة . فى هذا المثال يكون الغرض من الوصف هو تقديم وتفهم جيد لعدد كل من الذكور والاناث، والبنية العمرية للمجتمع المحلى، ونسب الأسر التى تقع داخل شرائح الدخل المختلفة.

فى هذا المثال، تمثل كل خاصية من الخصائص الثلاث، متغيرا. بمعنى أن يكون لكل خاصية هوية معينة يمكن أن تكتسب قيماً مختلفة لأفراد المجتمع المحلى الذى تعيش بداخلة. فمتغير النوع تكون له قيمتان احتماليتان هما ذكر، وأنثى. كما قد يتراوح متغير السن أو العمر من يوم واحد للاطفال حديثى الولادة إلى مائة عام للشيوخ وكبار السن. أيضا قد يأخذ متغير الدخل قيماً كثيرة على مدى واسع ومتصل يبدأ بأقل دخل للأسره وينتهى بأعلى دخل لها. ويعتبر التباين فى القيم للخصائص أو المتغيرات المعطاة فى هذا المثال الإهتمام الأساسى للوصف الإحصائى وتعرف مجموعة القيم لكل متغير من المتغيرات الثلاثة إحصائيا بتوزيع المتغير The distribution of the variable. ففى حالة متغير النوع مثلا،يشير توزيع الأفراد المقيمين داخل هذا المجتمع المحلى تبعا للقيم المصنفه بشكل معين ودالة على الذكور والإناث إلى توزيع الذكور والإناث فى المجتمع. أو أنها توزيع الافراد وفقاً لمتغير النوع.

٢- الإحصاء الاستدلالى : يتألف الإحصاء الاستدلالى من مجموعة الأساليب الإحصائية التى يستخدمها الباحثون فى الاستدلال على خصائص المجتمع الأصلى من خلال المشاهدات التى يتم إجراؤها على عينة ممثلة لهذا المجتمع. فمثلا، يستخدم الإحصاء الاستدلالى فى وصف المجتمع الكبير من خلال استخدام الباحثين للمعلومات من عينات صغيرة الحجم نسبياً من هذا المجتمع. وتنهض معظم البحوث الاجتماعية على الإحصاء الاستدلالى نظرا لصعوبة دراسة المجتمع الأصلى وتكلفة البحث الباهظة ماليا، وفزيقيا.

لو افترضنا أنك أردت دَراسة مجتمعك الذى تعيش فية ويبلغ عدد أفراده ٤٠٠٠٠ نسمة. فهل فى مقدورك أن تسأل كل فرد من أفراد هذا المجتمع حول سنه ودخل أسرته والتعرف على نوعه . وكم من الوقت المبذول والجهد الفيزيقى الشاق المطلوبين للقيام بهذا البحث فضلاً عن التكلفة المالية. التى يتم إنفاقها على فريق

من الباحثين أو المعاونين معك.

لذلك من الأفضل أن يلجأ الباحث إلى اختيار عينة ممثلة من هذا المجتمع ولنفترض مثلاً أنها تضم ٢٥٠ فرداً. فى هذه الحالة يسهل على الباحث أن يجمع من أفراد العينة جميع المعلومات الخاصة بالسن والدخل والعمر. ويمكن من خلال هذه المعلومات الاستدلال على شكل التوزيعات للمتغيرات الثلاثة على مستوى المجتمع الأصلى. ولو كان اختيار العينة صحيحا والوسائل المستخدمة فى جمع البيانات ملائمة لطبيعة المجتمع وأهداف البحث، سوف يحصل الباحث على وصف دقيق لخصائص هذا المجتمع.

لما كان الوصف الرقمى للعينات والمجتمع الأصلى ذات هدف متماثل ، كان ضروريا على الباحث أن يحافظ على ما يميز بينهما عندما يشير إلى أى منهما. فمثلا من خلال الاستدلال على خصائص المجتمع الأصلى نقول إن القيم تكون ثابتة بمعنى أن عدد كل من الذكور والاناث داخل هذا المجتمع فى أى وقت وعند أى نقطة ستمثل قيمة ثابتة لاتتغير ولتكن مثلا ٣٩٪ للذكور مقابل ٦١٪ للاناث . هذا الوصف الرقمى للذكور أو الاناث داخل المجتمع الأصلى، يعرف احصائيا بالمعامل The parameter (Kurtz , 1983: 2) وتبدو أهمية الإحصاء فى وصف العينة الذى لايعطى قيماً ثابتة أو متماثلة فى القيمة إذا قام الباحث بسحب أكثر من عينة من نفس النقطة وفى وقت واحد فسوف يحصل الباحث على نسب مئوية متباينة للذكور من عينة لأخرى. فمثلا قد تعطى خصائص العينة الأولى أن نسبة الذكور ٣٨.٢٪ ، بينما تعطى العينة الثانية ٣٩.٨٪ وهكذا . وهنا يبدأ دور الأساليب الأحصائية فى تحديد متوسط للتباين فى تقديرات العينة. وهنا نشير إلى الاوصاف الرقمية المسحوبة من العينة بأنها احصائيات Statistics وتكون فى الوقت ذاتة تقديريةلمعاملات (بارامترات)المجتمع الأصلى Populatian parameters.

من خلال مناقشتنا لتعريف الإحصاء وأساليبها المستخدمة فى العلوم الاجتماعية ذكرنا كلمات ، مثل البيانات Data، والمتغير The Variable،المجتمع الاصلى، العينة فماذا نعنى بهذه الكلمات إحصائيا ؟

تعريف البيانات ومصادرها :

يقصد بكلمة البيانات فى الإحصاء الشكل الرقمى من المعلومات التى تمثل خاصية أو ظاهرة ما. فمثلا فى العلوم السلوكية والنفسية ، لو كان إهتمام الباحث النفسى بدراسة مستوى القلق أو التوتر لدى جماعة من الأفراد، فان البيانات التى تمثل خاصية القلق تقع فى شكل قيم على مقياس نفسى للقلق. ويصبح تعريف الإحصاء هنا، مجموعة الأساليب المستخدمة فى وصف مستوى القلق لهذه الجماعة. وتستخدم البيانات فى وصف دور الإحصاء فى البحث الاجتماعى.

مصادر البيانات :

تتعدد مصادر البيانات فقد تكون مصادر تاريخية من واقع السجلات والملفات، وقد يكون المصدر من الميدان من خلال أساليب متنوعة يستخدمها الباحث فى جمع البيانات وفيما يلى نبذة عن كل مصدر:

المصدر التاريخى:

يشتمل هذا المصدر على جميع البيانات الإحصائية المدونةفى سجلات رسمية عن فترات زمنية ماضية ومحفوظة فى المؤسسات والهيئات (كالجهاز المركزى للتعبئة العامة والاحصاء، السجل المدنى ..... إلخ، والمنظمات الدولية مثل منظمة العمل الدولية ILO، والبرنامج الانمائى للامم المتحدة (UNDP) ...... إلخ.

فاذا أردنا، على سبيل المثال ، معرفة عدد سكان ريف وحضر مصر خلال الاعوام ١٩٧٦، ١٩٨٦، ١٩٩٦ . فيمكن معرفة ذلك من خلال البيانات المدونة فى التعدادات العامة للسكان والصادرة عن الجهاز المركزى للتعبئة العامة والإحصاء خلال السنوات المذكورة.

المصدر الميدانى:

يعتمد هذا المصدر على البيانات التى يقوم الباحث بجمعها من الميدان ، مستخدماً فى ذلك أسلوب أواكثر من أساليب جمع البيانات، وتعتبر صحيفة الاستبانة أكثر الأساليب شيوعاً فى الاستخدام فى المسوح الاجتماعية وتضم هذه

الصحيفة قائمة من الأسئلة (المفتوحة أو مغلقة النهاية) تغطى جوانب الظاهرة موضوع الدراسة وتطبق هذه الأداة من خلال المقابلة الشخصية او هاتفياً او بإرسالها بالبريد. ولسهولة التحليل الإحصائى ، يفضل ان تشتمل الاستبانةعلى اسئلة مغلقة النهاية أى تأتى الاستجابات لكل سؤال محددة. وكنموذج للأسئلة مغلقة النهاية ، نورد فيما يلى عدداً من الأسئلة التى تضمنتها صحيفة الاستبانة فى دراسة قامت بها المؤلفة تحت عنوان: النمو الحضرى والمدن الجديدة فى المجتمع القطرى (١٩٩٣).

أولاً: بيـانات عامـة :

فيما يلى بعض البيانات الأولية والمطلوب وضع علامة(√)أمام الإجابة الملائمة:

١ - النوع:

(١) ذكر.......................................( )

(٢) أنثى ....( )

٢ - الجنسية:

(١) قطرى.......................................( )

(٢) عربى ....( )

(٣) أجنبى ....( )

٣ - الحالة التعليمية:

(١) دون المتوسط ....( )

(٢) مؤهل متوسط ....( )

(٣) جامعى أو ما يعادله ....( )

(٤) ماجستير/ دكتوراه ....( )

٤- السـن :

(١) أقل من ٢٠ سنة ....( )

(٢) ٢٠ -٣٠ ....( )

(٣) ٣٠ - ٤٠ .... (   )
(٤) ٤٠ - ٥٠ .... (   )
(٥) ٥٠ - ٦٠ .... (   )
(٦) ٦٠ سنة فأكثر .... (   )

٥ - الحالة الاجتماعية:

(١) لم يسبق له الزواج .... (   )
(٢) متزوج .... (   )
(٣) مطلق .... (   )
(٤) أرمل .... (   )

٦ - قبل سن الثامنة عشر فى أى منطقة كنت تعيش مرحلتى الطفولة والصبا؟

(١) منطقة بدوية .... (   )
(٢) منطقة ريفية .... (   )
(٣) منطق حضرية .... (   )

٧ - مواعيد العمل الحالى:

(١) فترة صباحية فقط .... (   )
(٢) ورديات .... (   )
(٣) فترة صباحية وأخرى مسائية يومياً .... (   )

٨ - الكادر الوظيفى:

(١) فئة كبار الموظفين .... (   )
(٢) فئة الموظفين المتوسطين .... (   )

٩- مدة الخدمة بالشركة أو المؤسسات : بالسنوات:

(١) أقل من سنتين .... (   )
(٢) ٢-٤ .... (   )
(٣) ٤-٦ .... (   )

(٤) ٦-٨ .... (   )

(٥) ١٠ .... (   )

(٦) ١٠ سنوات فأكثر .... (   )

فيما يلي مجموعة من العبارات والمطلوب منك أن تقرأ كل عبارة على حدة وأن تحدد موقفك الشخصي منها، فإذا كانت تنطبق عليك تماماً فضع علامة (√) تحت كلمة أوافق بشدة أمام العبارة. أما إذا كانت تنطبق عليك على وجه التقريب فضع علامة (√) تحت كلمة أوافق إلى حد ما أمام العبارة. أما اذا كانت لاتنطبق عليك تماماً فضع علامة (√) تحت كلمة أعارض بشدة أمام العبارة. أما إذا كانت لاتنطبق إلى حد ما فضع علامة (√) تحت كلمة أعارض إلى حد ما أمام العبارة. وإذا كنت لا تستطيع تحديد موقفك فضع علامة (√) تحت كلمة لا أعرف أمام العبارة لا.

العبارات	أوافق بشدة	أوافق إلى حد ما	لا أعرف	أعارض إلى حد ما	أعارض بشدة
٣٢ - أنا على استعداد لتقديم أي جهد في سبيل أن تصبح مدينة أمسيعيد أجمل مدن قطر .					
٣٣ - دائماً أتحدث بحب مع أصدقائي ومعارفي عن مدينة أمسيعيد .					
٣٤ - أشعر بالسعادة لإقامتي وعملي داخل مدينة أمسيعيد.					
٣٥ - أتمنى الإقامة الدائمة في مدينة أمسيعيد.					
٣٦ - أقبل على الفور بالإلتحاق بعمل مماثل في مدينة الدوحة، إذ أتيح لي ذلك .					

العبارات	أوافق بشدة	أوافق إلى حد ما	لا أعرف	أعارض إلى حد ما	أعارض بشدة
٣٧ - أسهم بالجهد فى تنظيم وإقامة الحفلات التى تقيمها الأندية الإجتماعية دون التقيد بالعضوية لنادى معين.					
٣٨ - دائماً أشارك جيرانى مناقشة قضايا تهم مدينة أمسيعيد.					
٣٩ - أحرص دائماً على حضور الندوات وقراءة الموضوعات التى تدور حول تطوير مدينة أمسيعيد.					
٤٠ - اهتمامى بعملى يفوق بكثير اهتمامى بقضايا مدينة أمسيعيد.					
٤١ - لايهمنى إلاّ تحقيق التقدم فى عملى والاستفادة بمزاياه سواء كان هذا العمل فى مدينة أمسيعيد أم غيرها.					
٤٢ - لولا الاندماج فى العمل لساعات طويلة كل يوم لأحسست بالملل من إيقاع الحياة فى مدينة أمسيعيد.					
٤٣ - أهم ما يربطنى بمدينة أمسيعيد هدوئها والعلاقات الطيبة والتآلف بين المقيمين بداخلها.					

المجتمع الاصلي والعينة :

يشير المجتمع الأصلى Population إلى مجتمع البحث الذى يشتمل على جميع المفردات ويمكن سحب عينات بحثية منه. وذلك لصعوبة إجراء البحوث على جميع مفردات المجتمع الأصلى لاسيما كبير الحجم. لذا فإن معظم البحوث الاجتماعية تعتمد على المسوح بالعينة. وتشير العينة Sample إلى شريحة ممثلة للمجتمع الأصلى أى تشتمل على جميع خصائصه. وتوجد طرق عديدة لسحب العينات (*) من مجتمع البحث ، فعلى سبيل المثال يتم اختيارالعينة العشوائية إما بالطريقة البسيطة، او باستخدام قوائم الأرقام العشوائية، او بطريقة منتظمة او طبقية سواء نسبية أو غير نسبية ...إلخ. ونعنى بالعينة العشوائية Random Sample إعطاء فرص متساوية لجميع مفردات المجتمع لكى يكونوا ضمن مفردات العينة المختارة Subjects.

المتغير :

يعرف المتغير بأنه خاصية ما تأخذ قيماً مختلفة بين الأفراد داخل المجموعة قيد الدراسة. من أمثلة المتغير، خاصية الطول أو النوع أو الذكاء ، والاتجاهات حيث تتباين فى قيمها بين الأفراد داخل الجماعة الواحدة عند مرحلة معينة كالمرحله التعليمية الاعدادية أو الثانوية مثلا.

من جهه أخرى نجد خصائص أخرى تتصف بالثبات ويطلق عليها ثوابت Constants. اى تكون الخاصية ذاتها متكررة وموجودة. داخل كل فرد من أفراد جماعة ما . مثال ذلك مجموعة الطلاب فى الصف الثالث الاعدادى فإذا كان هؤلاء الطلاب مختلفين فى النوع والطول والذكاء(متغيرات) ، فإنهم متماثلون ويشتركون فى خاصية واحدة هى المرحلة التعليمية التى تعرف بأنها خاصية ثابتة. فالثابت The Constant هو خاصية ما تأخذ القيمة ذاتها لكل أفراد المجموعة قيد البحث.

(*) لمزيد من التفصيل حول المجتمع والعينة وطرق اختيارها انظر إعتماد علام ويسري رسلان، أساسيات الإحصاء الاجتماعي، دارقطري بن الفجاءة للنشر والتوزيع الدوحة – قطر ١٩٩١، ص ص ٢٧٠ – ٢٨٨.

تصنيف المتغيرات:

يتم تصنيف المتغيرات بعده طرق هى:

١ - تصنيف المتغيرات الكمية إلى متغيرات متصلة Continuous variables ، ومتغيرات متقطعة Discrete Variables . وسوف يتم شرحها فى الفصل الثانى .

أيضاً تنقسم المتغيرات إلى متغيرات معتمدة Dependent Variables ومتغيرات مستقلة Independent Variables

تعرف المتغيرات المستقلة بأنها المتغيرات التى يستطيع الباحث أن يتحكم فيها والتى يمكن تأويلها ومناقشتها تبعا للهدف البحثى. على سبيل المثال، لو كان الهدف من بحث اجتماعى هو معرفة تأثير إدمان السيدات للمخدرات على الصحة الانجابية لهن. فى هذا البحث يتحكم الباحث فى الجرعة من حيث الكمية والنوع ثم يلاحظ التغيرات الصحية على أفراد العينة. فى هذه الحاله تمثل الجرعة المخدره متغيراً مستقلاً والصحة الإنجابية متغيراً معتمداً وبالنسبة للبحوث الاجتماعية، تتعدد المتغيرات المستقلة التى تفسر حدوث ظاهرة ما. مثال ذلك العوامل التى تؤدى الى انحراف الاحداث قد تكون التفكك الأسرى ، مستوى تعليم الوالدين ، دخل الأسرة .... الخ .

يعرف المتغير المعتمد بأنه تابع للمتغير المستقل أى كلما تتغير قيم المتغير المستقل تتغير تبعا لذلك القيم المناظرة للمتغير المعتمد. ويستطيع الباحث أن يتعرف من خلال التغير فى قيم المتغير التابع كيفية ارتباطه بالمتغير المستقل .

مراحل البحث الإحصائى :

قسم الان جراهام Alan Graham (١٩٩٤) مراحل البحث الإحصائى إلى أربع مراحل أساسية لكل منها أساليب احصائية محددة . والمراحل هى :

مراحل البحث الإحصائى	الأدوات والاساليب الإحصائية
١ – صياغة وتوجيه الأسئلة	
٢ – تجميع البيانات	– اختيار العينة – تصميم صحيفة الاستبانة – تطبيق الأداة وجمع البيانات من المبحوثين .
٣ – تحليل البيانات	– حساب النسب المئوية – حساب المتوسطات – رسم منحنيات وأشكال بيانية
٤ – تفسير نتائج البحث	– صياغة التصورات المستقبلية على ضوء النتائج التى تم التوصل اليها . – اختبار العلاقة الارتباطية بين المتغيرات .

التعرف على خصائص واستخدام الاله الحاسبة Calculator فى الإحصاء :

لم يعد استخدام الحسابات اليدوية ملائما للعمليات الإحصائية المطلوبة فى البحوث الاجتماعية حتى مع أبسط الوسائل الإحصائية. لهذا يعتبر اكتساب الطالب والباحث مهارة استخدام الآله الحاسبة المحمولة أمراً ضرورياً كخطوة أولى على طريق استخدام الحاسب الآلى فى العمليات الإحصائية المتقدمة لاسيما فى مجال التطبيق لأساليب الإحصاء الاستدلالى.

تلعب الآله الحاسبة دوراً حساساً وايجابياً للغاية فى توفير الوقت وسرعة التحليل الإحصائى والتعامل بدقة عالية مع الحسابات التى قد لا تنتهى فى المعالجات الإحصائية. كما أصبحت الآله الحاسبة رخيصة الثمن ومتوفرة فى كل مكان. ولم يعد أمام الطالب أو الباحث الا أن يتعرف على النوعية الملائمة للآله

الحاسبة من حيث السعة والقدرة على التخزين والبرمجة للعمليات الحسابية المركبة. وتحمل الآله الحاسبة شاشه لقراءه العمليات الحسابية خلال إجرائها ثم التعرف على النتائج لكل عملية من خلال رموز إحصائية معينة مكتوبة على كل مفتاح من مفاتيح الآله والمرتبة أسفل الشاشة.

وعندما يبدأ الطالب فى تغذيه الآله الحاسبة بالأرقام واستخدام العمليات الحسابية الملائمة للهدف المنشود من جانبه ، يمكن أن يحصل على النتيجة من خلال الضغط على أحد المفاتيح التى تحمل كل منها العلامات الإحصائية الآتية:

مفتاح $\overline{X}$	- (سَ) عند حساب المتوسط الحسابى
$\sum X$	- (مجـ س) عند جمع قيم المتغير (س) كلها
$\sum X^2$	- (مجـ س^2)عند جمع مربع قيم (س) كلها
η	- عند طلب معرفة القيم التى تم ادخالها للعملية الحسابية
$X\sigma_\eta$	- (س سجما (ن)) وأحيانا تكتب فى بعض الآلات الحاسبة (σ_η) وتعنى حساب الانحراف المعيارى The Standard devation.
r	- (ر) حساب معامل الارتباط بين المتغيرات
a , b	- (أ ، ب) حساب معاملات الانحدار البسيط The Linear Regression Coefficients هذا بالإضافة إلى وجود مفاتيح اخرى للجذور والتربيع والعمليات الحسابية من جمع وضرب وطرح وقسمة.

وظائف الإحصاء فى البحوث الاجتماعية

١- تتدخل الإحصاء فى إفهام الباحث نوعية البيانات المتاحة واستقراء ماتعكسه من مؤشرات تتعلق بظاهرة معينة قيد البحث. كما تفيد الباحث فى إعداد جداول صحيحة إحصائيا للبيانات وجوده تنظيمها بما يسهل على القارىء لها من الوقوف على حال الظاهرة. وعلاقاتها بالظواهرة الأخري. ولذا يجب على الباحث أن يحدد بدقة الأساليب الملائمة لنوعية البيانات التى جمعها وأن يضع خطة للتحليل الإحصائى.

٢ - تتدخل الإحصاء فى مرحله التحليل للبيانات كمرحلة هامة من مراحل البحث الاجتماعى. ويكون الاستفاده بها أكثر عند بداية هذه المرحلة وبعد أن يتم سحب العينة البحثية من المجتمع الأصلى.

ومما يجدر التأكيد والإشارة إليه، أن الأساليب الإحصائية لاتضيف إلى أو تكون بديلا عن البيانات التى تم جمعها . فإذا افتقد الباحث القدرة على جمع البيانات الصحيحة حول الظاهرة التى يهتم بدراستها، فلن يكون استخدام الأساليب الإحصائية وسيلة لتحسن هذه البيانات أو تقويتها. أيضا اذا لم يتوفر لدى الباحث مهارة كافيه فى مجال الإحصاء تمكنة من استخدام أنسب أساليبها ملاءمة لهدف البحث سوف تكون النتائج التى يسفر عنها مفككة ويصعب مناقشتها والدفاع بقوة عما تشير إليه من قضايا مستقبلية أو راهنة تتعلق بموضوع البحث.

فى هذا الصدد، يقع الباحثون فى أخطاء عندما يقومون بجمع البيانات الخام غير الكافية وغير الشاملة لجوانب الظاهرة قيد الدراسة ، فإنه يصعب استخدام معاملات الارتباطات، كما تكون قياسات الصدق والثبات غير دقيقه . من هنا تتضح أهمية الالمام الجيد للباحث الاجتماعى بالأساليب الإحصائية بحيث يستطيع أن يختار من بينها ما يخدم أهداف البحث ، ويستطيع فى الوقت ذاته فهم ما تعكسة هذه الاساليب من نتائج (*).

تلعب الإحصاء دوراً حيوياً فى إعطاء الخاصية أو الظاهرة المقاسة معنى ومن ثم إضفاء مزيد من الفهم الواضح حولها من خلال التقديرات الإحصائية

(*) سوف يكون هذا الموضوع محور إهتمامنا في مؤلف لاحق إن شاء الله.

للبيانات الكمية والكيفية للظاهرة. ففى علم النفس الاجتماعى وعلم السلوك، قد يصعب إلى حد بعيد معرفة مستوى الإدراك والتوقعات والتوافق النفسى تجاه خاصية معينة دون استخدام الأساليب الإحصائية والمقاييس التى يجب فهمها وتوفر المهارة لدى الباحث فى تطبيقها. من أمثله المقاييس النفسية والسلوكية مقياس التوافق النفسى، مقياس التوتر، ومقياس الاغتراب النفسى.

تزداد اهمية الاستعانة بالاساليب الإحصائية خاصة الاستدلالية منها فى الابحاث العلمية والاجتماعية التى يتم اجراؤها على مجتمعات كبيره الحجم واحيانا تكون غير محدده أو معلومة الخصائص مثل الثروه البحرية أو المدفونه تحت باطن الأرض كالبترول والغازات وأثرها على المستوى الاقتصادى الاجتماعى لإفراد المجتمع ، على سبيل المثال. وقد يصعب تماماً دراسة المجتمع الكبير لأسباب مادية وفيزيقية حيث تتطلب فريقاً ضخماً من الباحثين يعملون لفترات زمنية طويلة وارتفاع التكلفة البحثية. لذلك يستخدم البحث الاجتماعى العينات التى تمثل المجتمع الأصلى. وتضم العينة مفردات تحمل خصائص هذا المجتمع. وتشير كلمة مفردة فى العينة إلى فرد أو حيوان أو أى شيء كالكتاب والمدرسة والسيارة ....إلخ.

تتدخل الإحصاء فى جانبين أساسيين: أولهما فى كيفية اختيار العينة، وثانيهما فى نوعية هذه العينة أو العينات التى تخدم أهداف البحث. ففى مجال اختيار العينة ، نجد أنها تعتمد على البيانات الإحصائية الدالة على صفات المجتمع الأصلى.

٣ – من تعريف الإحصاء ، نجد أن تلخيص البيانات تمثل الأهمية الثانية لها فى مجال البحث الاجتماعى. فإذا قام باحث اجتماعى بتوزيع حوالى مائتى استمارة استبانة وتضم كل منها حوالى ٧٠ سؤالا على أفراد مجتمع ما كمدينة القاهرة لاستطلاع الرأى حول برامج التليفزيون وحوراتها ، فى هذه الحالة سوف يتجمع لدى الباحث بيانات ضخمة لا يمكن له أن يتفهم مؤشراتها مالم يقم بتلخيصها باستخدام أساليب الإحصاء الوصفى من جدولة وتكرارت ورسومات بيانية ...إلخ. لأن الباحث إذا لم يكن لدية الخبرة بالمعالجة الإحصائية للبيانات، سوف يقدم تفسيرات ويحصل على نتائج غير دقيقة ولا تعكس بصدق الإتجاه العام للجمهور نحو البرامج التليفزيونية.

٤- تزداد أهمية استخدام الإحصاء فى تلخيص البيانات وتفسير النتائج خاصة عندما تتعدد المتغيرات التى يتم التعامل معها لخدمة أهداف البحث. فقد توجد متغيرات مستقلة عديدة نفسية واجتماعية وثقافية وتاريخية وسياسية ترتبط بالمتغير التابع فى الدراسة ، وتتطلب عملية التحليل للبيانات استخدام الانحدار المتعدد لتحديد الأهمية النسبية لهذه المتغيرات فى تفسير الظاهرة .

٥- تبدو أهمية استخدام الإحصاء فى البحوث الاجتماعية بمساهمتها الأساسية فى عملية تحويل المفهومات المجردة إلى متغيرات يمكن قياسها. وتشتمل هذه العملية على توصيف للعمليات التى يجب إعدادها واتخاذها لملاحظة أو لقياس ما يستدل به على المفهومات المجردة. ثم بعد تحديد المتغيرات وتعريفها إمبيريقيا، تبدأ عملية صياغه الفروض النوعية من حيث صلاحيتها فى تحديد ووصف نوع العلاقة بين المتغيرات. ونظراً لأن الفروض تمثل عبارات تربط بين متغيرات يمكن قياسها ، فمن الممكن بالتالى اختبار هذه الفروض من خلال الملاحظة والقياس. ويستخدم الباحثون وعلماء النظرية الاجتماعية الأساليب الإحصائيه المتنوعة فى استنباط ما يعرف بالنماذج العلمية Scientific models فى تحويل المقولات النظرية التى تؤلف نظرية اجتماعية ما إلى أبعاد يمكن قياسها مثال ذلك النماذج الرياضية والنماذج المستخدمة فى مجال الدراسات الاجتماعية لتنظيمات العمل . ويرتبط بهذا الدور الهام لاستخدامات الإحصاء فى البحث الاجتماعى، أهمية أخرى تتعلق باستخداماتها فى تقديم دليل أمبريقى ينهض على أساس رفض أو قبول أو تطوير نظرية مستخدمه فى مختلف العلوم السلوكيه والانسانية.

٦- يعتبر استخدام الأساليب الإحصائية هاماً فى تفسير نتائج البحث الاجتماعى وفى تقديم تفسيرات مدعمة إمبريقيا حولها هذا من شأنه أن يفسح المجال لدراسات مستقبلية تنطلق من استخلاصات صحيحة حول ظاهرة معينة، كما قد يفيد التفسير الإحصائى السليم لنتائج البحوث الاجتماعية الامبريقية فى عمليات التخطيط وصنع القرارات المتعلقة بالظاهرة قيد البحث ، وذلك من جانب المهتمين بها على المستويين الاكاديمى والعملى داخل تنظيمات العمل على اختلاف أنماطها داخل المجتمع.

٧- للأساليب الإحصائية دور بالغ الأهمية فى استخدام الحاسب الآلى The Computer فى البحث الاجتماعى اذ يتطلب استخدام الحاسب الآلى تعلم إحدى اللغات الخاصة والتى تمكن الباحث من التعامل معه على الوجه الصحيح . ومن أشهر الحزم الإحصائية المستخدمه فى العلوم الاجتماعية ما يعرف اختصارا بالأحرف الانجليزية SPSS وتوجد هذه الحزمة كبرنامج من برامج الحاسب الآلى، وتتحدد وظائفها فى تحليل البيانات فى سرعة وسهولة.

وتبرز أهميه استخدام الإحصاء مرحله ما قبل استخدام الحاسب الآلى فى تحويل وإعداد البيانات فى شكل أرقام واستخدام الترميز (التكويد) Coding حتى يسهل ادخالها لذاكره الحاسب الآلى لاسيما فى حالات تعدد المتغيرات واستخدام عدد من أساليب الإحصاء الاستدلالى التى تشتمل عليها حزمة SPSS.

المفاهيم الأساسية Key Concepts

الإحصـــاء :

مجموعة من الأساليب العلمية المستخدمة فى جمع البيانات وتنظيمها وتلخيصها وتحليلها بغرض الوصول إلى قوانين وقرارات .

الإحصاء الوصفى :

فرع من الإحصاء يختص بتلخيص توزيع متغير واحد وبقياس العلاقة بين متغيرين أو أكثر .

الإحصاء الاستدلالى :

فرع من الاحصاء يختص بعمل تصميمات للمجتمع الأصلى من خلال مايتم سحبه من عينات ممثله له .

العينـــة :

شريحة مختارة بعناية من المجتمع الأصلى ويتم الاختيار بعدة طرق ، وعلى الباحث أن يستخدم الطريقة الأكثر ملاءمة لأهداف بحثه وفى ضوء البيانات الإحصائية المتاحة عن خصائص المجتمع الأصلى .

المتغـــير :

أية خاصية يمكن أن تتغير قيمها من مفردة لأخرى داخل العينة البحثية . (كالطول ، الوزن ، العمر ، الحالة الزواجية) .

الثابت :

أية خاصية يفترض احتفاظها بقيمة ثابتة لجميع الافراد داخل مجموعة ما تحت الدراسة .

تمـــــارين

١ – أكمل العبارات الآتية :

(أ) ان مجموعة الأساليب الفنية المستخدمة من جانب علماء العلوم الاجتمـــاعية بقصد تنظيم ومناقشـــة البيانات بهـــدف اخـــتيار النظـــريات والاجـــابة على أســـئلة البـــحث ، تعـــد تعـــريفـــا لـ ....

(ب) تعتبر الاحصاء وتطبيق الأساليب الإحصائية من الأمور الحيوية من أجل .... البحث .

(جـ) يعرف .... بأنه خاصية قد تكتسب قيماً مختلفة من حالة إلى حالة أخرى .

(د) على الباحث الاجتماعى الذى يريد تلخيص ووصف توزيع ما أن يستخدم الأساليب الإحصائية لأحد فرعى الاحصاء المسمى ....

(هـ) إن الإحصاء .... يستخدمه الباحث عندما يريد فهم العلاقات بين متغيرين أو أكثر.

(و) يمثل الإحصاء الاستدلالى مجموعة الأساليب الإحصائية التى يستخدمها الباحثون عندما يرغبون فى التعميم من .... إلى ....

(ز) عندما تنشر صحيفة المساء المصرية استفتاء يشير إلى أن ٦٥٪ من المصريين يشاهدون كرة القدم ، فإنها استعانت بأساليب الإحصاء ....

٢ – فيما يلى عدد من الاختيارات إحداها يمثل الاجابة الصحيحة على كل سؤال من الأسئلة الآتية . ضع علامة (√) أمام الاختيار الذى تراه إجابة صحيحة على السؤال أعلاه .

ان قطر القمر بالأميال يمثل :

(أ) مجتمع أصلى .

(ب) ثابت .

(جـ) احصائى .

(د) مَعْلَم a parameter .

٣ - أجرى مسح اجتماعى على عينة عشوائية قوامها ٥٠٠ من الشباب فى حى صغير بمدينة القاهرة بهدف معرفة اتجاهاتهم نحو سياسة الحكومة الحالية مقارنة بالحكومة السابقة فيما يختص بالإهتمام بقضايا الشباب . فأى اختيار من الاختيارات الآتية يعتبر المتغير التابع :

(أ) حجم الحى الصغير فى مدينة القاهرة .

(ب) عدد المبحوثين فى العينة .

(جـ) الاتجاهات نحو سياسة الحكومة المصرية .

(د) الاختيارات الثلاثة السابقة ليس من بينها المتغير .

٤ - فى محاولة لتقدير سن الطالبات فى قسم الاجتماع بكلية بنات عين شمس ، قام أستاذ الإحصاء بأخذ متوسط العمر للطالبات داخل قاعة المحاضرات . فهل يمثل هذا المتوسط .

(أ) الإحصاء .

(ب) العينة .

(جـ) المجتمع الأصلى .

(د) مَعْلَم a parameter .

٥ - فى السؤال السابق ، تمثل الطالبات داخل قاعة المحاضرات بالنسبة لأستاذ الإحصاء .

(أ) إحصاء .

(ب) عينة .

(جـ) مجتمع أصلى .

(د) معْلَمْ .

٦ - هل الإحصاء الاستدلالى :

(أ) هو الإحصاء الوصفى .

(ب) يستخدم بعض أساليب الإحصاء الوصفى .

(جـ) تتيح للباحث التعميم للمجتمع الأصلى من خلال العينة .

(د) الاختياران الثانى والثالث السابقين معاً .

٧ - أراد أستاذ الإحصاء أن يقارن بين أسلوبين للتدريس لمادة الإحصاء فى فصلين دراسيين مختلفين . واعتمد فى المقارنة على الدرجات التى تحصل عليها الطالبات فى كل فصل منهما فى الاختبار النهائى لمادة الإحصاء . كم عدد المتغيرات التى سيقوم الأستاذ بإختبارها ؟ وما هو المتغير المستقل والمتغير المعتمد فى هذه الدراسة .

٨ - أجرى مسح قومى على عينة عشوائية قوامها ١٦٠٠ وحدة معيشية فى حى شبرا بمدينة القاهرة بهدف التعرف على الموافقة أو عدم الموافقة من جانب الأفراد على سياسة الحكومة المصرية فى تخفيض أسعار السلع الإستهلاكية . وكشفت نتائج المسح أن ٥٥٪ من الأفراد يوافقون على سياسة الحكومة . فى هذا المسح ماهو المجتمع الأصلى ؟ وماهى العينة ؟ وما نوع الأسلوب الإحصائى (استدلالى أم وصفى) الذى تم استخدامه فى هذا المسح ؟ .

٩ - فى عام ١٩٩٥ ، أشارت التقارير الإحصائية الرسمية إلى أن قيمة الوسيط للدخل السنوى للأسرة المصرية التى تضم أربعة أفراد يبلغ

٨٥٠٠ جنيهاً مصرياً . فى هذا المثال ، ماهو المجتمع الأصلى ؟ وماهى العينة ؟ وهل يعتبر هذا المثال نموذجاً للإحصاء الوصفى أو للإحصاء الاستدلالى ؟

١٠ – أراد فريق علماء الاجتماع السياسى أن يعرف الانتماءات الجزبية لطلاب جامعة القاهرة وذلك من خلال اجراء المقابلات معهم أثناء فترة تسجيلهم للعام الدراسى الجديد . وكشفت المقابلة عن أن ٤٥ ٪ من الطلاب ينتمون للحزب الوطنى ، ٣٠ ٪ لحزب الوفد ، ٢٥ ٪ لحزب العمل . فى هذا المثال ماهى العينة ؟ وماهو المجتمع الأصلى ، وأى الأساليب الاحصائية (وصفية أو استدلالية) ينتمى إليها هذا المثال ؟

* * * * *



الفصل الثانى

مستويات القياس والعرض الجدولي للبيانات

مقدمـــــة :

١ – المتغير المتصل .

٢ – المتغير المتقطع .

٣ – مستويات القياس .

٤ – التوزيع التكرارى .

٥ – الجداول التكرارية للبيانات الوصفية (الكيفية) .

٦ – الجداول التكرارية للبيانات الكمية .

٧ – الجداول المزدوجة .

٨ – الجدول الانتشارى .



الفصل الثانى

مستويات القياس والعرض الجدولى للبيانات

مقـــدمة :

من الشائع فى مجال البحوث الاجتماعية توافر نخبة من البيانات الإحصائية التى يحصل عليها الباحث باستخدام أدوات جمع البيانات المناسبة وعادة تتمثل تلك البيانات فى شكل أرقام تعتبر قياساً للمتغيرات تحت الدراسة ولما كانت تلك الأرقام تفتقر إلى الترتيب والتصنيف يطلق عليها البيانات الأولية أو البيانات الخام Row Data .

وتعرف البيانات الاحصائية إنها كمية من المعلومات على هيئة أرقام وإن تلك الأرقام أما أن تكون أرقاماً صحيحة Integers مثل ١٠ ، ٢٠ ، ٣٠وهكذا أو تكون أرقاماً عشرية أو حقيقية Real Numbers مثل ٨,٥ ، ١٠,٢٥ ، ١٥,٥ وهكذا .. ويتوقف حجم البيانات الخام على حجم المجتمع الأصلى ، فكلما إزداد حجم هذا المجتمع يتوقع مزيدا من الأرقام غير المرئية والتى يصعب مع كثرتها وعدم تصنيفها تفهم أو قياس متغير أو أكثر تحت الدراسة ، ومن ثم كان من الضرورى أن يقوم الباحث بتصنيف وتبويب تلك البيانات بالشكل أو بالأسلوب الذى يخدم جيداً هدف البحث من دراسة المتغيرات أو استنباط نوعية العلاقات أو المعلومات الهامة التى تتعلق بتلك المتغيرات فمثلاً لو أجرينا اختباراً لقياس القدرات لعدد (١٢٠) طالباً ، فإنه يلزم استخدام (١٢٠) رقماً مناظراً بواقع رقم لكل طالب. فلو افترضنا أن عدد الطلبة يصل إلى خمسة أو ستة أضعاف هذا العدد . فمن المؤكد أننا سوف نواجه صعوبات كثيرة فى قياس القدرة ما لم نستخدم أداة أحصائية أو أكثر لتنظيم البيانات الخام بهذا الحجم الكبير . ولعل أبسط الطرق الإحصائية لتنظيم وتلخيص البيانات هى طرق التوزيع التكرارى Frequency Distribution.

وقبل أن نتناول التوزيع التكرارى والوسائل الضمنية له فى تصنيف وتبويب البيانات – نرى أهمية الوقوف على نوعية البيانات الإحصائية من منظور مستويات القياس الإحصائى نظراً لأهمية تلك البيانات – فى التمييز بين حالة وأخرى من الحالات التى يكون عليها المتغير تحت الدراسة . هذا فضلاً عن أن اهتمامنا بتصنيف البيانات الإحصائية وفقاً لمستويات القياس الإحصائى يرجع إلى أن المتغيرات التى تقاس كمياً تنقسم من حيث قيمتها العددية إلى نوعين هامين هما :

المتغير المتصل *Continuous Variable*

لما كان التعريف العام للمتغير Variable هو ظاهرة أو صفة تختلف قيمها باختلاف الحالات ، فإن المتغير يكون متصلاً ، عندما يأخذ أى قيمة متدرجة على المقياس المستخدم . مثال ذلك قياس درجات الحرارة باستخدام الترمومتر . فالمتغير يأخذ قيمة ما بين رقمين صحيحين بمعنى أن المتغير يمكن أن يأخذ أى قيمة بين ٣٦ درجة ، ٣٧ درجة (٣٦,١ ، ٣٦,٢ ... إلخ) .

المتغير المتقطع *Discrete Variable* :

عندما يأخذ المتغير قيماً محددة يطلق عليه المتغير المتقطع ويكون المتغير المتقطع هو الذى يحتوى مداه على عدد محدود من القيم أو يحتوى على عدد لانهائى من القيم ولكن لكل منها قيمة محددة يمكن عدها أو ترتيبها فى نهاية الأمر . فعدد الأولاد أو الأفراد فى الأسرة لابد أن يكون رقماً صحيحاً مثل ١ ، ٢ ، ٣ ، ٤ ... وهكذا . ومن أمثال المتغيرات المتقطعة ، النوع Sex ، الحالة الزواجية Marital Status ، عدد أيام الانتاج فى أحد المصانع ، عدد حوادث السيارات ، وهكذا .

مستويات القياس :

تشتمل مستويات القياس على أربع مستويات هى :

وصفية Nominal ، ترتيبية Ordinal وأيضاً مقاييس فئوية والنسبة Interval and Ratio Scales .

يقصد بالقياس - كمفهوم واسع - إنه عملية تعبير عن الخصائص والمشاهدات بشكل كمى ووفقاً لقاعدة محددة . وعندما نستخدم المقياس ، فإننا لا نجد غضاضة فى اختيار نسق من المعادلات الرياضية التى تتفق مع تلك الخاصية أو الخصائص قيد البحث . وعامة يمكن القول أن ما تحظى به فروع العلم من نماذج متعددة فى الرياضيات والاقتصاد وغيرها من فروع العلوم الاجتماعية تعتمد فى بنيتها الأساسية على المقاييس .

ففى علم الاجتماع وعلم النفس الاجتماعى تتصف المتغيرات بالتباين والتعدد بشكل يصعب معه أن نختار نسقاً رياضياً مناسباً يخدم أهداف البحث الامبريقى . لأن الفرد بتركيبه النفسى يتصف بالتعقيد واختلاف مستويات العلاقة بينه وبين المحيطين به من أفراد أو بيئات .

ولعل أبسط أمثلة القياس نجدها فى الاختبارات المدرسية التى يتقدم لها الطلاب فى مختلف مراحل حياتهم الدراسية حيث ترتبط الدرجة التى يحصل عليها كل طالب فى اختبار ما على مدى معرفته بالمادة التى يدرسها خلال فترة دراسية معينة وكلما كانت درجة الطالب التى حصل عليها مثلاً فى مادة الكيمياء عالية دل ذلك على معرفة أكثر أو تحصيل أكبر لدى الطالب من هذه المادة . ومن هذا المثال البسيط نجد أن خاصية التحصيل تعبر عنها الدرجة Score التى حصل عليها الطالب من الاختبار .

وتعتبر المقاييس التى تقيس المتغير التابع Dependent Variable واحدة من أكثر المقاييس أهمية عند إيجاد الطرق الإحصائية الملائمة التى تستخدم فى تحليل بيانات دراسة امبريقية معينة .

أيضاً توجد بعض المقاييس التى يمكن استخدامها فى قياس ظاهرة معينة بدقة عالية أو متناهية مثال ذلك المقاييس التى تستخدم فى قياس الأطوال والأوزان. من جهة أخرى توجد بعض المقاييس التى تفتقر إلى الدقة العالية وإن كانت تحقق قدراً من الدلالة منها على سبيل المثال مقاييس مستويات القلق النفسى عند الأفراد .

ومن ثم تنقسم أنواع المقاييس وفق ترتيبها وما تتسم به من خصائص إلى

مقاييس تصنيفية ، مقاييس ترتيبية ، مقاييس فئوية ، وأيضاً مقاييس نسبه .

المقاييس التصنيفية :

يعتبر التصنيف أبسط العمليات الأساسية فى أى فرع من فروع العلم فالتصنيف هو تجميع للمفردات أو العناصر أو المعلومات المتشابهة إلى حد كبير والمتماثلة فى خصائصها مع بعضها فى مجموعة أو مصنف Category وذلك بهدف المقارنة بين المجموعات المختلفة على أساس الخواص . مثال ذلك إذا قمنا بتصنيف عدد من الأفراد إلى مجموعات وفق خاصية العقيدة Religion (مسلم ، مسيحى ، يهودى) . وقد نقوم أيضاً بعمل تصنيف آخر للنزعات السياسية للفئات الدينية الثلاث وهكذا .

ولابد من استخدام التصنيف كعملية أساسية تعتمد عليها المقاييس الأعلى كأساس لها أيضاً فى العلوم الاجتماعية . من ذلك لانبالغ بالقول إن التصنيف يعتبر المستوى الأول فى القياس .

خصائص المقياس التصنيفى :

فى المثال السابق ، نجد أننا لم نهتم بالتمييز بين الفئات الدينية الثلاث على أساس الأهمية مثلاً فلم نقل أن المسلم أهم من المسيحى أو أن المسيحى أهم من اليهودى ، فقط ينصب المقياس على تصنيف وفق الديانة وتمثل تلك الخاصية الأولى للمقياس التصنيفى والتى يمكن أن نحددها فى عدم اتصاف المقياس بالترتيب المنطقى .

كذلك نلحظ من المثال السابق ، عدم وجود أى تداخل على أساس الديانة فالمجموعة كاملة تضم أفراداً متماثلين فى نوع الديانة . ومن ثم لاتكون المفردة فى أكثر من مجموعة .

والخاصية الثالثة التى تتصف بها المقاييس التصنيفية فنجدها فى مجال العلاقات بين المفردات أو المقادير فى العلوم الرياضية على سبيل المثال . يتصف المقياس بخاصية الانتقالية Transitivity ويقصد بها أنه إذا كانت هناك علاقة معينة بين متغيرين ، بحيث أنها تتحقق من (أ) نحو (ب) فأنه من الضرورى أن

تتحقق أيضاً من المتغير (ب) نحو المتغير (أ) (Blalock, 1972, p. 15, 16, Hinkle; Wiersma and Jurs, 1979, p. 6)

المقياس الترتيبى

فى المثال السابق ، فضلاً عن تصنيف الأفراد إلى ثلاثة مذاهب دينية يمكن أن ترتب تلك المجموعات الثلاث وفقاً لأهميتها أو لما تمتلكه كل منها من خاصية أو سمات معينة مشتركة . وقد نجد مثالاً أقرب للفهم فى الرياضيات عندما تميز بين المقدارين (أ) ، (ب) فنقول أن (أ) أكبر من (ب) ونأخذ الشكل الرياضى التالى :

أ > ب

وقد تكون أ > ب ولكن مقدار الفرق فى القيمة الدالة على التمييز بين أ ، ب ليس من خصائص المقياس الترتيبى . ومن ثم فإن المقياس الترتيبى ذا مستوى أعلى من المقياس التصنيفى فى قياس الظواهر أو الخواص . وتعتبر خاصية التمييز باستخدام علامات (>) أو (<) الخاصية الثانية إذا أخذنا فى الاعتبار خاصية التصنيف وفق الترتيب .

وفى العلوم الاجتماعية نجد مثالاً لخاصية الترتيب دون الالتزام بالفروق عندما نصنف الأسر وفقاً للمكانة الاجتماعية – الاقتصادية Socioeconomic Status إلى طبقة عليا ، ووسطى وأيضاً إلى طبقة دنيا .

وتشير الخاصية الثالثة إلى عدم تكرار نفس المفردة فى أكثر من مجموعة كما هو الحال فى المقياس التصنيفى .

والخاصية الرابعة فهى الانتقالية . فلو فرضنا أن أ > ب وأن ب > جـ فيمكن القول أن أ > جـ وهذه خاصية أخرى يتشابه فيها هذا المقياس مع المقياس التصنيفى ولكن من المنظور الترتيبى .

ويجدر التنوية إلى ضرورة ملاحظة أن المستوى الترتيبى للمقياس لا يهتم بالفروق بين العناصر أو الخواص. ومن ثم لانستطيع أن نستخدم مع هذا المقياس العمليات الحسابية كالطرح والقسمة والضرب والجمع كما أننا لا يمكن استخدامها

أيضاً مع المقياس التصنيفى .

ولتوضيح هذه الملاحظة الهامة لو افترضنا أن هناك أربع نقاط على متصل ويرمز لها بالأحرف (أ ، ب ، جـ ، د) ويفارق مسافات معينة تقع النقطتان ب ، جـ بين النقطتين (أ) ، (د) كما فى الشكل التالى :

متصل

أ　　ب　　جـ　　د

فباستخدام المقياس الترتيبى يمكن كتابة العلاقة التالية (اتجاهياً)

$\overline{\text{أد}} = \overline{\text{أب}} + \overline{\text{ب جـ}} + \overline{\text{جـ د}}$ (Blalock, 1972, p. 16)

ولكن لا يمكن إطلاقاً معرفة أطوال المسافات الأربعة المبينة فى العلاقة السابقة . مثال ذلك الترتيب المستخدم فى مقاييس الاتجاهات الذى يبدأ بالموافقة بشدة وينتهى بعدم الموافقة .

المقياس الفـئوى :

من خصائص المقياس الفئوى Interval Scale بالإضافة للخصائص التى ذكرناها فى المقياسين السابقين ، توحيد نوع وحدة القياس فلايمكن أن نقيس الفرق بين درجتين من الحرارة إحداهما بالفهرنهايت والأخرى بالدرجة المئوية بل يكون الفرق بين درجتين حراريتين مثل ٣٨ درجة مئوية، ٣٠ درجة مئوية أى من نفس جنس وحدة القياس .

من جهة أخرى ، إذا قلنا أنه توجد وحدات قياسية للقياس الفئوى ، ففى العلوم الاجتماعية قد يتعذر تحقيق ذلك ، فمثلاً لا توجد وحدات قياسية أو معيارية لقياس الذكاء ، السلطة ، الهيبة الاجتماعية والتى نجدها متكررة دائماً فى الموضوعات الاجتماعية والنفسية المختلفة .

والخاصية الثانية للمقاييس الفئوية إمكانية استخدام العمليات الحسابية المختلفة من جمع وطرح وضرب وقسمة للدرجات فى عمليات تحليل البيانات . فمثلاً يمكن إضافة دخل الزوجة إلى دخل الزوج أو إلى دخل باقى أفراد الأسرة .

الخاصية الثالثة للمقياس الفئوى أنه يهتم بخاصية تساوى الفروق بين المستويات المختلفة مثال ذلك تقسيم الدرجة الواحدة على مقياس الحرارة (الترمومتر) إلى خمسة أقسام يمثل كل جزء منها ٠,٢ من الدرجة . ويطلق على هذا النوع من المقاييس مقياس الفئات المتساوية Equal intervals .

القياس النسبى

يعتبر القياس النسبى Ratio من أرقى مستويات القياس ويشتمل على جميع الخصائص السابقة . فضلا عن وجود الصفر المطلق الذى يعنى غياب الخاصية . والقياس النسبى ليس محور اهتمامنا فى البحوث الاجتماعية .

التوزيع التكرارى

يعرف التوزيع التكرارى Frequency Distribution بأنه عملية ترتيب الأرقام فى صورة تعطى عدد مرات حدوث الرقم أو الصفة أو مانسميها بالتكرارات. هذا إلى جانب أن التوزيع التكرارى ينظم البيانات نمطياً كما يعطى الباحث بعض الدلالات عن طبيعة البيانات التى بين يديه . من جهة أخرى فإن التوزيع التكرارى لايعطى الباحث أى معلومة حول تلك البيانات بل يمثل مقدمة لاجراء التحليل الاحصائى واختبار الفروض .

خصائص التوزيع التكرارى :

توجد أربع خصائص تحدد المفهوم السابق للتوزيع التكرارى وهى :

الوضع المركزى Central Location :

ونعنى بها قيمة محددة تتوسط باقى قيم التوزيع . وتستخدم المتوسطات كمقاييس إحصائية فى تقدير تلك القيمة والمقاييس للمتوسطات هى المتوسط الحسابى ، المنوال ، والوسيط ، وتعرف بمقاييس النزعة المركزية والتى سنعرض لها فى الفصل الرابع . ونظراً لاختلاف طريقة الحساب من مقياس إلى آخر فى حساب المتوسطات ، فإن القيمة الوسطى سوف تختلف قيمتها تبعاً لذلك .

التشتت أو التباين

ويقصد به درجة انتشار أو تشتت القراءات المختلفة للظاهرة حول قيمتها الوسطى . وأنه كلما إزداد تجمع المفردات حول تلك القيمة الوسطى قل تشتتها

وبالتالى يقال للتكرارات التجريبية عن تلك الظاهرة أنها أكثر تجانساً .

الالتـــــواء :

وهى خاصية تعنى ابتعاد التوزيع التكرارى عن التماثل ويمكن قياس الالتواء Skewness بعدة طرق من بينها المنوال والوسيط ، الربيعين وأيضاً استخدام العزوم . وقد يطلق على التوزيع التكرارى أنه موجب الالتواء Positively Skewed إذا كانت معظم التكرارات متجمعة عند القيم الكبرى ناحية اليمين . أما إذا كانت معظم التكرارات متجمعة عند القيم الصغرى من ناحية اليسار من المنحنى التكرارى فيطلق على التوزيع أنه سالب الالتواء Negatively Skewed ويمكن حساب الالتواء باستخدام معادلة كارل بيرسون من خلال العلاقة الآتية :

المتوسط الحسابى – المنوال = ٣ (المتوسط الحسابى – الوسيط)

وقد اقترح بيرسون مقياسين لحساب الالتواء هما :

(أ) معامل بيرسون الأول للالتواء = $\frac{\text{المتوسط الحسابى – المنوال}}{\text{الانحراف المعيارى}}$

(ب) معامل بيرسون الثانى للالتواء = $\frac{\text{٣(المتوسط الحسابى – الوسيط)}}{\text{الانحراف المعيارى}}$

ويجدر التنوية أن القسمة على الانحراف المعيارى تهدف إلى تحويل معامل الالتواء إلى مقياس نسبى يمكن به مقارنة الالتواء فى التوزيعات المختلفة .

مـــثال (١)

كان المتوسط الحسابى والمنوال والانحراف المعيارى لمجموعة من القيم كما يلى :

المتوسط = ٣٣,١٧٥

المنوال = ٣٢,٠٦

الانحراف المعيارى = ٥,٥٦

والمطلوب حساب الالتواء باستخدام معامل بيرسون الأول للالتواء :

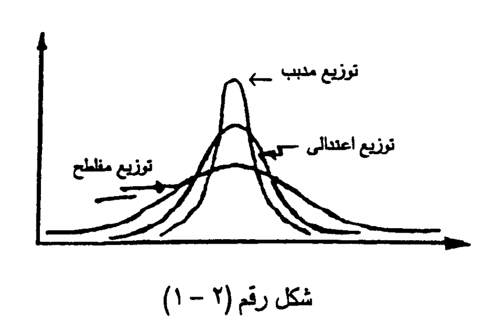
$$\text{معامل الالتواء} = \frac{\text{المتوسط الحسابى} - \text{المنوال}}{\text{الانحراف المعيارى}}$$

$$= \frac{٣٣,١٧٥ - ٣٢,٠٦}{٥,٥٦} = ٠,٢$$

ويلاحظ أن الإلتواء موجب الإشارة أى متجهاً نحو اليمين ويمكن باستخدام خاصية الإلتواء تحديد مدى تماثل التوزيع التكرارى . فالتوزيعات التكرارية تكون متماثلة عندما يساوى معامل الإلتواء صفراً .

التفلطـــح :

تعتبر أهمية هذه الخاصية فى أنها تحدد مدى اختلاف التوزيع التكرارى للظاهرة عن التماثل للتوزيع الاعتدالى . بالإضافة إلى خاصية الإلتواء . فالتفلطح Kurtosis ينبأ عما إذا كان للتوزيع التكرارى للقيم قمة علياً حادة أو قمة عريضة مسطحة . ففى الشكل رقم (٢ – ١) نجد لدينا ثلاثة منحنيات تكرارية أوسطها توزيع اعتدالى Normal Distribution أما المنحنيان الآخران فالأول له قمة مدببة تعلو قمة المنحنى الاعتدالى وهذا المنحنى له خاصية التدبب ويطلق عليها التوزيع الإنحنائى المنحنى المدبب ، أما المنحنى الثالث والذى تقل قمته عن قمة التوزيع الإعتدالى وتأخذ القمة شكلاً أكثر إستوائية وتفلطح تبعاً لتشتت القيم التى يشملها التوزيع ويطلق عليه منحنى تكرارى مفلطح . ومن ثم يمكن أن نميز بين التوزيعين نسبياً بالتوزيع الإعتدالى كما يلى :



شكل رقم (٢ – ١)

أ – **التوزيع المدبب** : هو التوزيع الذى تكون تكرارته أكبر من التكرارات فى التوزيع الاعتدالى كما تزداد فيه خاصية تجمع التكرارات بالقرب من الفئات الوسطى .

ب – **التوزيع المفلطح** : هو التوزيع الذى تقل تكراراته المركزية عن التوزيع الاعتدالى كما تنتشر تكراراته على مدى أكبر حول الفئات الوسطى ويمكن قياس التدبب والتفلطح على أساس الفرق بين مدى ارتفاع قمة المنحنى الاعتدالى عن المحور الأفقى وكلا من قمتى التوزيع المدبب والمفلطح على التوالى .

حساب معامل التفلطح :

يمكن حساب قيمة معامل التفلطح للتوزيعات التكرارية من المعادلة التالية :

$$\text{معامل التفلطح} = \frac{\text{نصف الفرق بين الربيعين الأعلى والأدنى}}{\text{الفرق بين المئينين التسعون والعاشر}}$$

$$= \frac{\text{الانحراف الربيعى}}{\text{المئين التسعون} - \text{المئين العاشر}}$$

مـــثال (٢)

تم حساب القيم للمقاييس الأربعة فى المعادلة السابقة لقياس معامل التفلطح لإحدى التوزيعات التكرارية . فكانت تلك القيم كالآتى :

الربيع الأعلى = ٧٢,٤٥

الربيع الأدنى = ٥٦,٥٥

المئين التسعون = ٧٩,٤٥

المئين العاشر = ٤٤,٧٥

والمطلوب حساب معامل التفلطح لتلك التوزيعات ؟

الحـــل :

$$\text{معامل التفلطح} = \frac{\frac{٧٢,٤٥ - ٥٦,٥٥}{٢}}{٧٩,٤٥ - ٤٤,٧٥} = \frac{٧,٩٥}{٣٤,٧٠} = ٠,٢٢٩$$

وباستـخدام الخصـائص الأربع السابقـة للتـوزيع التكرارى يمكن أن نقسم التوزيعات التكرارية بشكل عام إلى نوعين أساسيين هما :

أ – توزيعات تكرارية اعتدالية

ب – توزيعات تكرارية غير متماثلة Asymmetrical distribution :

وتعرف بأنها التوزيعات التى تتناقص أو تزداد فيها التكرارات للقيم بصورة غيـر اعتـدالية أو غيـر منتظمـة على جانبى المحور الرأسى المقام عند منتـصف التوزيع والذى يقطع المحور السينى أو المحور الأفقى فى التمثيل البيانى .

العـرض الجـدولى للبيـانـات :

أوضحنا فيما سبق أهمية التوزيع التكرارى كوسيلة لتصنيف البيانات الخام حول الظاهرة المطلوب دراستـها وتجـميع بيانات يتم تحـويلها إلى بيانـات رقـمية تمهيداً لتحليلها إحصائياً . وقلنا أن البيانات الرقمية قد تكون كثيرة وذات قيم مختلفة سواء كانت متقاربة فى قيمها أو متباينة . ومن ثم لايجد الباحث مفراً من ضرورة تصنيف تلك البيانات ومحاولة وضعها فى شكل جداول تمكن من عرضها بصـورة تلخص مـعالمها وتساعـد على استـخـلاص النتـائج منهـا وهذا مـا نعنيـه إحصائياً بعملية التبويب Tabulation .

ويتم التبويب للبيانات عادة على أساس كمى أو كيفى (نوعى) أو جغرافى أو زمنى . كمـا يمكن أن يتم التـبـويب بأسلوب المزج بين هذه الأسس . وتصنف المتـغيـرات إلى متـغيـرات كـمـية (مـثل الدخل ، العمر ، الوزن ، الطول ، درجـة الحرارة) ، أو متغيرات كيفية أو أسمية (مثل الحالة الزواجية ، النوع ، المهنة والجنسيـة) . فبالنسبة للمـتغير الكمى ، تحمل القيمـة معنى كمياً ويتم ترتيب

مفردات البحث حسب الخاصية موضوع الدراسة ، كالعمر مثلاً ، فى هذه الحالة ، يتم الترتيب من الأكبر سناً إلى الأصغر سناً أما فى حالة المتغير الكيفى ، فإن القيمة لا تشير إلى مقدار الخاصية ، بل تعبر إما عن وجود تلك الخاصية أو غيابها. وقد يعطى الباحث قيماً رقمية لصفات المتغير كأن يعطى رقم (١) ليدل على أن المبحوث اعزب ورقم (٢) للمتزوج فإن هذه الأرقام لا تشير إلى مقدار الخاصية أو أن رقم (٢) أكبر من رقم (١) . بل تستخدم هذه الأرقام لغرض تصنيف صفات المتغير وليس لها اى دلالة رقمية.

الجداول التكرارية للبيانات الوصفية (الكيفية)

تشير البيانات الوصفية إلى صفات نوعية توضح عناصر الظاهرة موضوع الدراسة مثل الديانة ، المهنة ، منطقة السكن ، الحالة الزواجية والجنسية . ويتم الحصول على تلك البيانات باستخدام أى أسلوب من الأساليب السابقة لجمع البيانات .

(أ) الجداول التكرارية البسيطة

مثال (٣)

حصل أحد الباحثين على البيانات التالية والتى تتعلق بالجنسية لمجموعة من طلاب جامعة عين شمس . والمطلوب عمل جدول تكرارى بسيط يوضح توزيع الطلاب حسب الجنسية .

مصرى	أجنبى	عربى	مصرى	أجنبى	عربى
مصرى	أجنبى	عربى	مصرى	عربى	مصرى
عربى	عربى	أجنبى	مصرى	أجنبى	مصرى
مصرى	مصرى	أجنبى	مصرى	مصرى	عربى
مصرى	مصرى	عربى	عربى	عربى	أجنبى
مصرى	مصرى	مصرى	عربى	أجنبى	مصرى

خطوات حل المثال :

١ – لتنظيم هذه البيانات فى جدول توزيع تكرارى يلزم عمل جدول تفريغ (جدول رقم ٢ – ١) يشتمل على ثلاثة أعمدة :

(أ) يتضمن العمود الأول عناصر الظاهرة وهى مصرى ، عربى ، وأجنبى. ويكون عنوان هذا العمود الجنسية .

(ب) يشتمل العمود الثانى على العلامات حيث يتم تسجيل المشاهدات على شكل خطوط رأسية مائلة تساوى فى عددها عدد التكرارات للصفة الواحدة . فإذا تكررت الصفة مرة واحدة يسجل خط رأسى مائل «/» ، وإذا ما وصلت عدد العلامات إلى أربع تكتب أربع خطوط رأسية مائلة «////» ، ثم يكتب الخط الخامس أفقياً ليقطع الأربعة خطوط المائلة ويعرف هذا الشكل بالحزمة «~~////~~» وعددها خمس تكرارات . وهكذا يتم تسجيل جميع القراءات فى أشكال خطوط رأسية مائلة حتى أربعة تكرارات أو حزمة وتكرارتها .

(جـ) أما العمود الثالث فيكون بعنوان التكرارات ويقصد بها عدد عناصر الظاهرة أمام كل صفة من الصفات الموجودة فى العمود الأول .

جدول رقم (٢ – ١)

تفريغ البيانات المتعلقة بجنسية الطلاب

الجنسية	العلامات	التكرارات
مصرى	~~////~~ ~~////~~ ~~////~~ //	١٧
عربى	~~////~~ ~~////~~ /	١١
أجنبى	~~////~~ ///	٨
المجموع		٣٦

٢ – من الجدول السابق يتم عمل جدول التوزيع التكرارى ويشتمل على عمودين : أولهما بعنوان اسم المتغير (الجنسية) ؛ وثانيهما يتضمن التكرارات المناظرة لكل صفة من صفات المتغير . أى أن جدول التوزيع التكرارى يشتمل

فقط على العمودين الأول والثالث من جدول تفريغ البيانات الخام . وذلك على النحو التالى :

جدول رقم (٢ – ٢)

التوزيع التكرارى للطلاب حسب الجنسية

الجنسية	ك
مصرى	١٧
عربى	١١
أجنبى	٨
المجموع	٣٦

(ب) الجداول التكرارية المزدوجة للبيانات الوصفية

تشتمل الجداول التكرارية المزدوجة على متغيرين كالتوزيع التكرارى لافراد العينة حسب الاصول الريفية الحضرية والجنسية ، أو حسب النوع والمهنة ، أو الجنسية والنوع .

مثال (٤) :

عند دراسة الجنسية والحالة الزواجية لعينة عشوائية تم سحبها من احدى الشركات الاستثمارية بمدينة العاشر من رمضان ، كانت النتائج على النحو التالى:

مصريون			عرب			أجانب		
متزوج	متزوج	لم يسبق	متزوج	لم يسبق	متزوج	متزوج	لم يسبق	مطلق
متزوج	لم يسبق	له الزواج	ارمل	له الزواج	متزوج	ارمل	له الزواج	متزوج
لم يسبق	له الزواج	لم يسبق	مطلق	متزوج	ارمل	متزوج	متزوج	
له الزواج	متزوج	له الزواج	متزوج	ارمل	متزوج	متزوج	متزوج	
لم يسبق	ارمل	ارمل	متزوج	ارمل		متزوج	لم يسبق	
له الزواج	متزوج	مطلق	متزوج	متزوج		ارمل	له الزواج	
ارمل	ارمل	متزوج	لم يسبق	لم يسبق		لم يسبق	متزوج	
مطلق		متزوج	له الزوج	له الزواج		له الزواج	مطلق	

والمطلوب عمل جدول تكرارى مزدوج لعرض هذه البيانات .

خطـوات الحـل :

١ - عمل جدول تفريغ مزدوج بحيث يشتمل العمود على صفات المتغير الأول الحالة الزواجية (لم يسبق له الزواج ، متزوج ، ارمل ، مطلق) . ويشتمل الصف على صفات المتغير الثانى الجنسية (مصريون ، عرب ، وأجانب) . وتشتمل كل صفة فى الصف على عمودين أولهما للعلامات وثانيهما للتكرارات المناظرة لعدد القراءات المسجلة بالعلامات أمام كل صفة فى العمود كما يتضح من الجدول رقم (٢-٣) .

جدول رقم (٢ - ٣)

تفريغ البيانات المتعلقة بالجنسية والحالة الزواجية

الجنسية / الحالة الزواجية	مصريون		عــرب		أجــانب		المجموع
	العلامات	ك	العلامات	ك	العلامات	ك	ك
لم يســبق له الزواج	~~////~~	٥	///	٣	///	٣	١١
متزوج	~~////~~ //	٧	~~////~~ ////	٩	~~////~~ ///	٨	٢٤
ارمل	////	٤	////	٤	//	٢	١٠
مطلق	//	٢	/	١	//	٢	٥
المجموع		١٨		١٧		١٥	٥٠

٢ - عـمل جـدول توزيـع تكرارى لأفـراد العينة حـسب الجنسـيـة والحـالة الزواجية من جدول التفريغ السابق على النحو التالى

جدول رقم (٢ - ٤)

التوزيع التكرارى لافراد العينة حسب الجنسية والحالة الزواجية

الحالة الزواجية / الجنسية	مصريون	عرب	أجانب	المجموع
لم يسبق له الزواج	٥	٣	٣	١١
متزوج	٧	٩	٨	٢٤
أرمل	٤	٤	٢	١٠
مطلق	٢	١	٢	٥
المجموع	١٨	١٧	١٥	٥٠

جدول رقم (٢ - ٥)

التوزيع التكرارى النسبى والمئوى لجنسية الطلاب

الجنسية	ك	التكرار النسبى	٪
مصرى	١٧	٠,٤٧	٤٧,٢
عربى	١١	٠,٣١	٣٠,٦
أجنبى	٨	٠,٢٢	٢٢,٢
المجموع	٣٦	١	١٠٠,٠

جدول رقم (٢ - ٦)

التوزيع التكرارى المئوى للطلاب حسب الجنسية والحالة الزواجية

الجنسية / الحالة الزواجية	مصريون		عرب		أجانب		المجموع	
	ك	٪	ك	٪	ك	٪	ك	٪
لم يسبق له الزواج	٥	٢٧,٨	٣	١٧,٦	٣	٢٠,٠	١١	٢٢,٠
متزوج	٧	٣٨,٩	٩	٥٢,٩	٨	٥٣,٣	٢٤	٤٨,٠
أرمل	٤	٢٢,٢	٤	٢٣,٥	٢	١٣,٣	١٠	٢٠,٠
مطلق	٢	١١,١	١	٥,٩	٢	١٣,٣	٥	١٠,٠
المجموع	١٨	٣٦,٠	١٧	٣٤,٠	١٥	٣٠,٠	٥٠	١٠٠,٠

الجداول التكرارية للبيانات الكمية (الرقمية)

إن الباحث أثناء قيامه بتصنيف البيانات الكمية له أن يختار الفئات التى يحددها لنفسه ولكن بشرط أن تسهل عدد الفئات ومداها من إدراك ما بين البيانات الإحصائية من علاقات وما لها من صفات ودلالات . ونعنى بالفئة تلك المجموعة الرقمية الجزئية داخل مدى البيانات التكرارية بحيث تشتمل على عدد من القيم المتقاربة . وبعد ذلك يقوم الباحث بحصر العدد الذى يقع داخل حدى كل فئة (الحد الأدنى والحد الأعلى) ويسمى تكراراً ويرمز له بالرمز (ك) . وبعد تكرار ذلك العمل لجميع البيانات داخل المدى المحدد لها بأعلى قيمة وأدنى قيمة من خلال ترتيب تنازلى أو تصاعدى فإن الباحث سيحصل فى النهاية على مايسمى بالجدول التكرارى والذى يتضمن عدداً من الفئات وتكرارتها والذى يصبح أساس دراسة الظاهرة قيد البحث . ويطلق على أبسط الطرق لتنظيم البيانات الإحصائية بالمصفوفة الرقمية أو العددية Array وهذه الطريقة تصلح إذا كانت القيم صغيرة. أما إذا كانت القيم بالآلاف ومضاعفاتها فيصعب تماماً استخدام المصفوفة ومن هنا تبرز أهمية استخدام الجدول التكرارى وأيضاً المنحنيات التكرارية Frequency Curves .

الجداول التكرارية البسيطة :

ذكرنا أن الجداول التكرارية Frequency Tables تعتبر إحدى وسائل عرض البيانات فى شكل يسهل معه التحليل والوصول للنتائج التى تتطلبها الدراسات . حيث أن تجميع البيانات فى شكل فئات ذات تكرارت تقلل من حجم توزيع البيانات ومن ثم يسهل استيعاب القيم الواردة بالجدول .

ولتوضيح طريقة عمل الجدول التكرارى نأخذ المثال التالى :

مثال (٥) :

قام مجموعة من الأساتذة والمعيدين باستلام أوراق امتحان مادة الاحصاء لعدد ١٨٠ طالباً وطالبة بالفرقة الثانية قسم الاجتماع . وتبعاً لترتيب استلام هذه الأوراق من مراقبى اللجان ، تم رصد مجموع الدرجات لكل طالب من الدرجة العظمى للامتحان وهى (١٠٠) . على النحو التالى . والمطلوب عمل جدول توزيع تكرارى لدرجات الطلاب .

الحـــل :

هذه البيانات الخام يصعب تحديد أعلى الدرجات وأقلها . فإذا أردنا معرفة ذلك لزم أعادة تنظيم تلك الدرجات كأن نرتبها ترتيباً تنازلياً بمعنى أن نبدأ بالدرجة الأعلى فى القيمة ثم تليها الأقل منها مباشرة فى القيمة وهكذا حتى تصل الى أقل درجة .

٦٨	٥٢	٦٩	٥١	٤٣	٣٦	٤٤	٣٥	٥٤	٥٧	٥٥	٥٦
٥٥	٥٤	٥٤	٥٣	٣٣	٤٨	٣٢	٤٧	٤٧	٥٧	٤٨	٥٦
٦٥	٥٧	٦٤	٤٩	٥١	٥٦	٥٠	٤٨	٥٣	٥٦	٥٢	٥٥
٤٢	٤٩	٤١	٤٨	٥٠	٢٤	٤٩	٢٥	٥٣	٥٥	٥٢	٥٦
٦٤	٦٣	٦٣	٦٤	٥٤	٤٥	٥٣	٤٦	٥٠	٤٠	٤٩	٤١
٤٥	٥٤	٤٤	٥٥	٦٣	٥٥	٦٢	٥٦	٥٠	٤٦	٤٩	٤٧
٥٦	٣٨	٥٥	٣٧	٦٨	٤٦	٦٧	٤٥	٦٥	٤٨	٦٤	٤٩
٥٩	٤٦	٥٨	٤٧	٥٧	٥٨	٥٦	٥٩	٦٠	٦٢	٥٥	٦٣
٥٦	٤٩	٥٥	٥٠	٤٣	٤٥	٤٢	٤٦	٥٣	٤٠	٥٢	٤١
٤٢	٣٣	٤١	٣٤	٥٦	٣٢	٥٥	٣٣	٤٠	٤٥	٣٩	٤٦
٢٨	٤٣	٣٧	٤٤	٥٤	٥٦	٥٣	٥٧	٥٧	٤٦	٥٦	٤٥
٥٠	٤٠	٤٩	٣٩	٤٧	٥٥	٤٦	٥٤	٣٩	٥٦	٣٨	٥٥
٣٧	٢٩	٣٦	٣٠	٣٧	٤٩	٣٦	٥٠	٣٦	٤٤	٣٥	٤٥
٤٢	٤٣	٤١	٤٢	٥٢	٤٧	٥١	٤٦	٦٣	٤٨	٦٢	٤٩
٥٣	٦٠	٥٢	٦١	٤٩	٥٥	٤٨	٥٦	٣٨	٤٨	٣٧	٤٧

أما الخطوة الثانية فتتمثل فيما يسمى بجداول التفريغ بأن يقوم الدارس بوضع خط رأسى مائل أمام كل رقم يتكرر بحيث يصبح عدد العلامات مساوياً لعدد مرات تكرار كل رقم . (انظر مثال ٣) وبهذا الشكل ينخفض حجم الجدول ويتحدد فى أعمدة ثلاثة ويراعى أن يأتى الترتيب فى الجدول تسلسلياً مبتدأ بأعلى رقم ومنتهياً بأقل رقم كما يتضح من جدول رقم (٢ – ٧) .

الترتيب التنازلى للدرجات :

٦٩	٦٢	٥٧	٥٦	٥٤	٥٢	٤٩	٤٨	٤٦	٤٣	٤٠	٣٦
٦٨	٦٢	٥٧	٥٥	٥٤	٥٢	٤٩	٤٨	٤٦	٤٣	٤٠	٣٦
٦٨	٦٢	٥٦	٥٥	٥٤	٥٢	٤٩	٤٧	٤٦	٤٣	٣٩	٣٦
٦٨	٦١	٥٦	٥٥	٥٤	٥١	٤٩	٤٧	٤٥	٤٢	٣٩	٣٥
٦٥	٦٧	٥٦	٥٥	٥٤	٥١	٤٩	٤٧	٤٥	٤٢	٣٩	٣٥
٦٥	٦٠	٥٦	٥٥	٥٣	٥١	٤٩	٤٧	٤٥	٤٢	٣٨	٣٤
٦٤	٥٩	٥٦	٥٥	٥٣	٥٠	٤٩	٤٧	٤٥	٤٢	٣٨	٣٣
٦٤	٥٩	٥٦	٥٥	٥٣	٥٠	٤٩	٤٧	٤٥	٤٢	٣٨	٣٣
٦٤	٥٩	٥٦	٥٥	٥٣	٥٠	٤٩	٤٧	٤٥	٤١	٣٨	٣٣
٦٤	٥٨	٥٦	٥٥	٥٣	٥٠	٤٨	٤٦	٤٥	٤١	٣٧	٣٢
٦٣	٥٨	٥٦	٥٥	٥٣	٥٠	٤٨	٤٦	٤٤	٤١	٣٧	٣٢
٦٣	٥٧	٥٦	٥٥	٥٣	٥٠	٤٨	٤٦	٤٤	٤١	٣٧	٣٠
٦٣	٥٧	٥٦	٥٥	٥٢	٥٠	٤٨	٤٦	٤٤	٤١	٣٧	٢٩
٦٣	٥٧	٥٦	٥٤	٥٣	٤٩	٤٨	٤٦	٤٤	٤٠	٣٧	٢٥
٦٣	٥٧	٥٦	٥٤	٥٢	٤٩	٤٨	٤٦	٤٣	٤٨	٣٦	٢٤

جدول رقم (٢ - ٧)

تفريغ درجات امتحان مادة الإحصاء

الدرجة	العلامات	ك	الدرجة	العلامات	ك
٦٩	/	١	٤٦	~~////~~ ////	٩
٦٨	//	٢	٤٥	// ~~////~~	٧
٦٧	/	١	٤٤	////	٤
٦٦	صفر	صفر	٤٣	////	٤
٦٥	//	٢	٤٢	~~////~~	٥
٦٤	////	٤	٤١	~~////~~	٥
٦٣	~~////~~	٥	٤٠	////	٤
٦٢	///	٣	٣٩	///	٣
٦١	/	١	٣٨	////	٤
٦٠	//	٢	٣٧	~~////~~	٥
٥٩	///	٣	٣٦	////	٤
٥٨	//	٢	٣٥	//	٢
٥٧	/ ~~////~~	٦	٣٤	/	١
٥٦	//// ~~////~~ ~~////~~	١٤	٣٣	///	٢
٥٥	// ~~////~~ ~~////~~	١٢	٣٢	//	٣
٥٤	// ~~////~~	٧	٣١	صفر	صفر
٥٣	// ~~////~~	٧	٣٠	/	١
٥٢	/ ~~////~~	٦	٢٩	/	١
٥١	///	٣	٢٨	صفر	صفر
٥٠	// ~~////~~	٧	٢٧	صفر	صفر
٤٩	/ ~~////~~ ~~////~~	١١	٢٦	صفر	صفر
٤٨	/// ~~////~~	٨	٢٥	/	١
٤٧	// ~~////~~	٧	٢٤	/	١

جدول رقم (٢ - ٨)

التوزيع التكرارى لدرجات امتحان الإحصاء

ف	ك
٢٠ -	١
٢٥ -	٢
٣٠ -	٧
٣٥ -	١٨
٤٠ -	٢٢
٤٥ -	٤٢
٥٠ -	٣٠
٥٥	٣٧
٦٠ -	١٥
٦٥ - ٧٠	٦

نلاحظ فى الجدول (٢ - ٧) إننا رتبنا الأرقام بفاصل واحد فقط إلا أنه فى بعض الحالات يمكن إعادة تنظيم الارقام على شكل مجموعات رقمية مسلسلة بفاصل أكبر من الواحد كأن يكون الفاصل مقداره الحسابى خمسة وهذا الفاصل يسمى طول الفئة Class Interval وهذا من شأنه يعطى مقارنة للبيانات بصورة أكثر سهولة .

ففى المثال السابق يمكن إعادة ترتيب مجاميع الأرقام لتصبح كما هو موضح فى الجدول رقم (٢ - ٨) بفاصل خمسة بين كل فئة . ونجد أننا أخذنا فئات متساوية المدى أو الطول وهو (٥) أو بمعنى آخر يمكن القول أن الفرق بين مركز كل فئة ومركز الفئة التى يليها يكون مساوياً ومقداره (٥) درجات . ويعرف مركز الفئة بأنه أما يساوى الحد الأدنى للفئة مضافاً إليه نصف طولها أو يساوى الحد الأعلى لتلك الفئة مطروحاً منه نصف طولها.

ولايشترط أن تكون جميع فئات الجدول التكرارى متساوية فقد تقتضى

الظروف أن نحدد فئات غير متساوية الطول كما هو الحال فى جداول توزيع الحيازة الزراعية أو التوزيع العمرى للسكان .

كذلك لايشترط فى تحديد الفئات أن تكون محددة القيمة لحديها الأدنى والأعلى فقد توجد فئات مفتوحة الحد الأعلى مثل جداول الأعمار مثال ذلك الأعمار فوق ٦٠ سنة . ووفقاً لنوعية الفئات وتقسيماتها بهذه الصورة تنقسم الجداول التكرارية إلى نوعين :

الجداول التكرارية المقفلة :

وهى الجداول التكرارية التى تكون حدود جميع الفئات معروفة تماماً سواء الحد الأعلى أو الحد الأدنى لكل منها وإن لم تكن مكتوبة صراحة . وتكون أطوال الفئات متساوية ويمكن بمعرفة طول الفئة تحديد الحد الأعلى للفئة الأخيرة .

الجداول التكرارية المفتوحة :

وهذا النوع من الجداول يشتمل على فئتين غير محددتى النهاية فالفئة الأولى ليس لها حد أدنى وأيضاً الفئة الأخيرة يكون الحد الأعلى لها غير معلوم . فإذا تحقق هذا الشرط للفئتين فى جدول تكرارى فيطلق عليه جدولاً تكرارياً مفتوح الطرفين . أما إذا تحقق ذلك فى فئة واحدة منهما فيطلق عليه جدولاً تكرارياً مفتوح الطرف الواحد الأدنى أو الأعلى . وفى هذا النوع من الجداول يستحيل استخدام المقاييس الإحصائية مثل المتوسط الحسابى والتباين .

أسس اختيار عدد الفئات :

يتوقف تحديد عدد الفئات على ظروف الظاهرة قيد البحث فإذا اختار الباحث عدداً قليلاً من الفئات فإن ذلك يعنى وجود عدد كبير من التكرارات داخل طول الفئة وبالتالى تكون احتمالية تباين قيمها كبيراً . ومن جهة أخرى لو اختار الباحث عدداً كبيراً من الفئات فقد يؤدى ذلك الى قرب الجدول التكرارى من البيانات الخام . هذا فضلاً عن وجود شرط أساسى عند اختيار الفئات يتحتم على الباحث ضرورة تحديد الفئات بصورة لاتجعل معها وجود فئات ذات تكرارات كثيرة بينما تخلو فئات أخرى من التكرارات . ومن ثم يتوقف اختيار عدد الفئات وطول كل فئة على خبرة الباحث . ولتيسير الأمر على الباحث ، فقد جرى العرف

بين الإحصائيين ألا يقل عدد الفئات عن ٥ فئات ولاتزيد إطلاقاً عن ١٥ فئة باستثناء البيانات الخاصة بالمسوح الاجتماعية لتعداد السكان . ومن المستحسن عموماً جعل الفئات متساوية الطول بقدر الإمكان .

ويمكن الحصول على طول الفئة الذى يناسب عدد القيم الكلى باستخدام المعادلة الرياضية الآتية وبالتالى يمكن تحديد عدد مناسب من الفئات .

$$\text{طول الفئة} = \frac{\text{المدى المطلق}}{(١ + ٣,٣٢٢ \times \text{لوغاريتم عدد القيم})}$$

$$\text{عدد الفئات} = \frac{\text{المدى المطلق}}{\text{طول الفئة}}$$

ويراعى ضرورة توافر الشروط التالية عند اختيار عدد الفئات :

١ - أن يتم اختيار عدد الفئات بحيث تغطى المدى الكلى للبيانات التكرارية .

٢ - أن يفترض الباحث توزيعاً متجانساً للقيم داخل كل طول فئة يتم اختيارها .

٣ - يتم اختيار عدد الفئات بحيث لا يكون أكبر من المناسب لضمان حسن التلخيص والتبويب وخشية إضاعة بعض معالم التوزيع نتيجة عملية إدماج بعض المفردات مع البعض الآخر .

ويمكن الحصول على طول الفئة بطريقة أيسر باستخدام المعادلة التالية :

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات المقترح}}$$

المدى = الفرق بين أعلى قيمة وأقل قيمة

تطبيق على مثال (٥) :

المدى = ٦٩ - ٢٤ = ٤٥

$\text{طول الفئة} = \frac{٤٥}{١٠} = ٤,٥$ درجة ٥ درجات تقريباً

طـــرق كـتـابـة الـفـــئـات :

(أ) فى حالة البيانات المتقطعة :

يقوم الباحث بتحديد الحدين الأدنى والأعلى لكل فئة بوضوح رقمياً فضلاً عن كتابتها على هذا النحو فى توزيعات عدد أفراد الأسرة :

١ – ٣ فرد إلى ثلاثة أفراد .

٤ – ٦ أربعة إلى ستة أفراد .

(ب) فى حالة البيانات المتصلة :

يقوم الباحث بتوضيح الحد الأعلى لكل فئة بينما يتم تحديد الحد الأدنى بها ضمنياً مع تجنب حدوث تداخل بين نهاية فئة وبداية الفئة التى تليها هكذا حتى لا يتكرر إزدواج لعدد من المفردات .

مثال : فى توزيع المفردات وفقاً للفئات العمرية تكتب على النحو التالى :

– ٢٠

– ٣٠

– ٤٠

كما يمكن كتابة الحدود الدنيا فقط للفئات على الشكل التالى :

١٠ –

٢٠ –

٣٠ –

جـدول الـتـوزيع الـتـكـرارى الـنـسـبـى :

يقصد بالتكرار النسبى للفئة ، كما سبق وأوضحنا ، هو تكرارها بالنسبة إلى التكرار الكلى لجميع الفئات أو بالأحرى هو حاصل قسمة قيمة تكرارات الفئة على مجموع قيم التكرارات لجميع الفئات الموجودة بالجدول وعادة يعبر عن هذا الناتج فى شكل نسبة مئوية أى يتم ضرب الناتج فى ١٠٠ للحصول على تلك النسبة (٪) ويطلق على التكرار فى هذه الحالة بالتكرار المئوى .

والغرض الأساسى من عمل جداول التكرار المئوى هو احتياج الباحث فى بعض الأحيان إلى معرفة نسبة كل تكرار إلى التكرار الكلى . فمثلاً عندما يريد الباحث معرفة أكثر الأسر الريفية إنجاباً للأطفال أو أعلى نسبة وفيات بين الأطفال

بسبب مرض معين ، أو أعلى الفئات الوظيفية تقاضياً للأجور والرواتب الشهرية فى شركة ما أو لمجموع شركات فى نفس المجال النوعى للصناعة .

مثال (٦) :

فيما يلى بيانات الأجور الأسبوعية لعدد (٢٥) عاملاً باليومية فى إحدى الشركات الصناعية ، وأرادت إدارة الشركة معرفة أكثر هؤلاء العمال انتظاماً ومواظبة فى العمل وذلك من خلال :

١ - أن أعلى الأجور تدل على الانتظام فى العمل حيث لايتقاضى عامل اليومية أجراً عن كل يوم غياب .

٢ - إن أقل أجر يدل على عدم الانتظام والمواظبة فى الحضور للعمل . وفى هذا المثال يمكن استخدام جداول التكرار النسبى والمئوى ، حيث أن النسبة المئوية لكل فئة تحدد أكثر الفئات وأقلها مواظبة وأيضاً الفئات متوسطة المواظبة فى العمل.

١٨	٢٤	٣٢	٢٤	١٦
٢٧	١٤	٢٥	٤٠	٣٦
١٢	٢٣	٣٤	٣٨	١٧
١٨	٣٨	٢٤	٤٢	٢٢
٢٧	١٦	٣٥	٢٠	٢٦

جدول رقم (٢ - ٩)

جدول تكرارى نسبى لاجور العمال

الفئات	التكرار	التكرار النسبى	التكرار المئوى
١٠ -	٢	٢٥/٢ = ٠,٠٨	٨٪
١٥ -	٥	٢٥/٥ = ٠. ٢	٢٠٪
٢٠ -	٦	٢٥/٦ = ٠,٢٤	٢٤٪
٢٥ -	٤	٢٥/٤ = ٠,١٦	١٦٪
٣٠ -	٤	٢٥/٢ = ٠,٠٨	٨٪
٣٥ -	٤	٢٥/٤ = ٠,١٦	١٦٪
٤٠ - ٤٥	٢	٢٥/٣ = ٠,٠٨	٨٪
مجموع	٢٥	مجموع نسبى = ١	مجموع مئوى = ١٠٠٪

من الجدول يتضح أن أكثر الفئات انتظاماً ومواظبة هى فئتا الأجر (٢٠-) ، (١٥-) وأقل الفئات انتظاماً ومواظبة هى فئات الأجر (٤٠-) ، (٣٠-) ، (١٠-) .

جداول التكرار التجمعى Cumulative Frequency Tables :

يستخدم هذا النوع من الجداول التكرارية إذا اراد باحث أن يعرف عدد المفردات التى تقل أو تزيد عن قيمة معينة وذلك نظراً لأن الجدول التكرارى المتجمع يسهل معرفة تكرارات كل فئة على حدة . ولعل من أهم مايتميز به الجدول التكرارى التجمعى هو السهولة فى الوصول الى التكرارات المطلوبة .

وهناك نوعان من التكرارات التجمعية وفقاً لترتيبها أما تصاعدياً فيطلق عليها الجدول التكرارى المتجمع الصاعد ، وقد ترتب التكرارات تنازلياً ويطلق عليها الجدول التكرارى المتجمع الهابط .

فلكى يحصل الباحث على قيم بعض المفردات التى تزيد فى القيمة عن قيمة معينة ، فيقوم الباحث بجمع القيم الأقل من الحد الأعلى للفئة المعينة . ومن ثم يقع مجموع تلك التكرارات فى مدى نفس الفئة . ويراعى أن تكون عملية جمع التكرارات الأقل متوالية فى ترتيب يبدأ من أحد طرفى الجدول حتى طرفه الآخر من أجل الحصول على المجموع الكلى للتكرارات .

ففى حالة التكرار المتجمع الصاعد يكون التجميع من أعلى إلى أسفل ويبدأ الباحث تسلسلياً ابتداءاً من الفئات الأقل حتى يصل إلى تكرارات الفئات العليا هذا ويلاحظ أن تكون التكرارات المتجمعة فى زيادة مستمرة حتى تصل إلى أكبر تكرار والذى يمثل الحد الأخير فى التسلسل . هذا ويطلق على عمود ترتيب الفئات الصاعدة أما (أقل من الحد الأعلى) أو (الحدود العليا للفئات) وكلاهما صحيح .

وفى حال التكرار المتجمع النازل ، يكون التكرار المتجمع للفئة الأخيرة مساوياً فى القيمة لتكرارها بالجدول . ويبدأ الباحث فى تجميع التكرارات من أسفل إلى أعلى مبتدءاً بالفئات الكبيرة تنازلياً حتى ينتهى بالفئات الأقل فالأقل حتى تكون أقل قيمة تكرارية هى الفئة الأخيرة بالجدول . وفى كلا النوعين من الجدول التجمعى ، يجمع الباحث تكرار كل فئة على مجموع التكرارات السابقة عنها مع

رصد المجموع الكلى أمام تلك الفئة ويطلق على عمود ترتيب الفئات تنازلياً (الحد الأدنى فأكثر) أو (الحدود السفلى للفئات) وكلاهما صحيح أيضاً .

مثال (٧) :

وضح نوعى الترتيب التجمعى لعدد الأفراد القاطنين فى إحدى المبانى السكنية وفقاً للفئات العمرية .

جدول رقم (٢ - ١٠)

جدول تكرارى متجمع صاعد وهابط

الجدول التكرارى		الجدول التكرارى المتجمع الصاعد		الجدول التكرارى المتجمع الهابط	
فئات (عمرية)	عدد الأفراد من القاطنين	الحدود العليا للفئات	التكرار المتجمع الصاعد	الحدود الدنيا للفئات	التكرار المتجمع الهابط
٢٠ -	١٤	أقل من ٢٥	١٤	٢٠ فأكثر	٧٠
٢٥ -	٨	أقل من ٣٠	٢٢	٢٥ فأكثر	٥٦
٣٠ -	٧	أقل من ٣٥	٢٩	٣٠ فأكثر	٤٨
٣٥ -	٥	أقل من ٤٠	٣٤	٣٥ فأكثر	٤١
٤٠ -	١٦	أقل من ٤٥	٥٠	٤٠ فأكثر	٣٦
٤٥ - ٥٠	٢٠	أقل من ٥٠	٧٠	٤٥ فأكثر	٢٠
مجـ	٧٠				

الجــداول المزدوجـــة

تستخدم فى تلخيص إزدواج القيم لمتغيرين . حيث يتم تبويب البيانات وفقاً لصفتين فى ترتيب صفوف وأعمدة بحيث تشمل الصفوف تكرارات الصفة الأولى بينما تشمل الأعمدة تكرارات المتغير الثانى . كما يمكن استخدامها أيضاً وبنفس التنظيم فى حصر عددى . مثال ذلك الجدول المزدوج التالى الذى يشتمل على حصر عدد الوفيات خلال عام فى الريف والحضر داخل محافظات الوجهين البحرى والقبلى وأيضاً فى المحافظات الحضرية كما يتضح من جدول رقم (٢-١١).

جدول رقم (٢-١١)

توزيع عدد الوفيات (بالألف) على مستوى الجمهورية (ريف وحضر)

الوفيات/المحافظات	عدد الوفيات (بالألف)		الجملة
	حضر	ريف	
الوجـــه البحـرى	٤٨	٥٢	١٠٠
الوجـــه القـــبلى	٥٠	٥٦	١٠٦
المحافظات الحضرية	٥٨	-	٥٨
المجموع	١٥٦	١٠٨	٢٦٤

أيضاً من أنواع الجداول المزدوجة ما يشتمل على ظاهرتين تتصفين بأكثر من نوعين ويراد معرفة العلاقة بينهما . ويطلق على هذا النوع الجداول التوافقية والمثال التالى يوضح جدول توافق Contingency Table تقديرات النجاح لعدد ١٠٠ طالباً وطالبة فى مادتى الإحصاء والاجتماع الصناعى فى قسم الاجتماع :

جـدول رقـــم (٢-١٢)

جـدول توافـقـى

مادة الإحصاء / مادة الإجتماع	ضعيف	مقبول	جيد	ممتاز	المجموع
ضـعيف	١١	٧	٥		٢٣
مقـــبول	٨	٣٠	١٠		٤٨
جـــــيد	٢	١٤	١٠	١	٢٧
ممـــتاز	-	-	١	١	٢
المجموع	٢١	٥١	١٦	٢	١٠٠

من أنواع الجـداول المزدوجـة أيضـاً ما يسـتـخـدم فى تلخـيص إزدواج القيم لمتـغـيرين تحت الدراسـة لمعرفـة درجـة ونوع العـلاقـة بينهـمـا . وهذا النوع شـائع الاسـتـخدام فى مـجـال الإحـصـاء لإيجاد مـعامل الارتبـاط لبيرسـون (ر) . والذى يستنبط منه مايعرف بالجداول الهامشية Marginal Tables .

ومن أمثلة جداول العلاقة المزدوجة ، تلك التى تستخدم فى إيجاد العلاقة بين الرضا عن العمل ومتوسط الأجر ، العلاقة بين الحالة التعليمية ومستوى دخل الأسرة ، العلاقة بين الوعى الانتخابى والحالة التعليمية ، العلاقة بين التغيب والاحساس بالاغتراب فى مكان العمل وغيرهم من الأمثلة .

خطوات عمل الجدول التكرارى المزدوج (العلاقة بين متغيرين) :

أ - حدد مدى التغير لكل متغير على حدة .

ب - قسم كل مدى منهما إلى عدد مناسب من الفئات .

جـ - قم بإعداد الجدول من أعمدة رأسية تتضمن تكرارات المتغير الأول مثلاً وليكن (ص) وأيضاً صفوف أفقية تتضمن تكرارات المتغير الثانى وليكن (س) . ويوضح المثال التالى جدولاً مزدوجاً يتضمن العلاقة بين عدد أيام الغياب (س) خلال عام مع متوسط الأجر لعينة مكونة من ١٠٠ عامل فى إحدى الشركات الصناعية (انظر جدول رقم ٢-١٣) .

جدول رقم (٢ - ١٣)

جدول مزدوج للعلاقة بين س ، ص

عدد أيام الغياب (س)	١٥ -	٢٥ -	٣٥ -	٤٥ -	٦٠ فأكثر	المجموع
			الدخل (ص)			
٤٠ -			١٠		١٠	٢٠
٨ -	١	٤	٤	٥		١٤
١٢٠ -	٤		٤			٨
١٦٠ -		٥	٦			١١
٢٠٠ -	٦	٤				١٠
٢٤٠ -	١	٨	٦	٢		١٧
٢٨٠ -	٨	٢				١٠
٣٢٠	٩	١				١٠
المجموع	٢٩	٢٤	٣٠	٧	١٠	١٠٠

كيفية استنباط الجدول الهامشى من الجدول المزدوج للعلاقة بين متغيرين :

فى هذا يقوم الباحث بعمل جدول هامشى لكل متغير تمهيداً لحساب معامل ارتباط بيرسون وذلك بأن يأخذ للمتغير الأول (ص) العمود الأول والعمود الأخير من الجدول المزدوج بكل التكرارات التى يتضمنها . ويقوم بإعداد جدول من صفوف وأعمدة على غرار الجدول المزدوج . ويكرر نفس العمل للمتغير الثانى (س) ولكن يأخذ الصف الأول العلوى والصف الأخير (المجموع) ، فمن المثال للعلاقة بين الأجر والتغيب عن العمل لعينة ١٠٠ عامل يمكن استنباط الجدول الهامشى بهده الطريقة :

جدول رقم (٢ - ١٤)

الجدول الهامشى للمتغير (ص)

المتغير ص	التكرارات
٤٠ -	١٠
٨٠ -	١٤
١٢٠ -	٨
١٦٠ -	١١
٢٠٠ -	١٠
٢٤٠ -	١٧
٢٨٠ -	١٠
٣٢٠ -	١٠
مجموع	١٠٠

جدول رقم (٢ - ١٥)

الجدول الهامشى للمتغير (س)

المتغير س	التكرارات
١٥ -	٢٩
٢٥ -	٢٤
٣٥ -	٣٠
٤٥ -	٧
٦٠ فأكثر	١٠
مجموع	١٠٠

الجدول الانتشارى

وهو نوع من الجداول المزدوجة يتضمن تكرارات لمتغيرين والمراد حساب العلاقة بينهما أيضاً إلا أن الفرق بين جدول العلاقات المزدوج السابق والجدول الانتشارى أن الباحث يقوم فى الأخير بوضع علامة واحدة تعبر عن كل قيم فى المتغير الأول وبالنسبة للمتغير الثانى يقوم بوضع علامة أخرى واحدة أيضاً إلا أنها تعبر عن قيمتين الأولى وهى الخاصة بالمتغير الأول والثانية الخاصة بالمتغير الثانى وخطوات عمل الجدول الانتشارى تتلخص فى الآتى :

١ - يقوم الباحث بتحويل قيم كل متغير إلى فئات وبالتالى يحصل على فئات للمتغير الأول وفئات للمتغير الثانى .

٢ - يوضح تدرج الفئات فى الجدول الانتشارى فى العمود الرأسى مثلاً للمتغير الأول وفى الصف الأفقى الأعلى للمتغير الثانى مع وجود عمود للمجموع بالنسبة للأول وصف للمجموعَ بالنسبة للمتغير الثانى .

٣ - يقوم الباحث بتفريغ كل قيمتين متناظرتين من قيم (س ، ص) لكل حالة والبحث عن الفئة التى يقع خلال طولها كل قيمة من القيمتين المتناظرتين . ويكرر ذلك لكل حالة من الحالات المعطاة للمتغيرين .

٤ - يقوم الباحث بوضع علامة مائلة (/) هكذا فى الخانة التى تلتقى فيها قيمتى (س، ص) المتناظرتين لكل حالة . وهذه العلامة تعبر عن نوع العلاقة بين المتغيرين . ففى النهاية نحصل على الجدول المزدوج الانتشارى حيث يكون اتجاه العلامات قطرياً ، أما أن يبدأ من أعلى الجدول يميناً ويتجه نحو اليسار إلى أسفل الجدول وفى هذه الحالة تكون العلاقة بين المتغيرين تامة موجبة .

وأما إذا كان اتجاه العلامات المائلة قطرياً متجهاً من أعلى الجدول من اليسار متجهاً إلى أسفله ناحية اليمين فإن العلاقة بين المتغيرين تكون تامة سالبة (-) .

مثال (٨) :

أوجد نوع العلاقة بين المتغيرين س ، ص وفقاً للحالات الأربع التالية باستخدام الجدول الانتشارى :

الحالة	س	ص
١	٣٨	٢٢
٢	٣٤	١٩
٣	٢٥	١٨
٤	٣٩	٢٤

الحل :

١ - يتم عمل الجدول الانتشارى بوضع قيم المتغير (س) فى فئات وأيضاً قيم المتغير (ص) فى فئات بحيث توضع فئات كل متغير فى الجدول ولنفرض أن فئات المتغير (ص) فى العمود الرأسى وفئات المتغير (س) فى الصف الأفقى الأعلى على هذا النحو :

جدول رقم (٢ - ١٦)

فئات س	٢٥ -	٣٥ -	المجموع
١٥ -	// أ	ب	
٢٠ -	جـ	// د	
المجموع			

٢ - الجدول أننا قمنا بتفريغ كل قيمتين متناظرتين للمتغيرين (س ، ص) فى كل حالة على حدة وفى خانة الإلتقاء يتم وضع علامة (/) ، ففى الحالة الأولى نجد أن لهما قيمتين من (س ، ص) هما (٣٨) ، (٢٢) . وحيث أن القيمة الأولى للمتغير (س) وهى (٣٨) تقع فى طول الفئة (٣٥-) ويناظرها فيه (ص) وهى القيمة (٢٢) والتى تقع فى طول الفئة (٢٠-) من فئات المتغير (ص) فالقيمتان يلتقيان رأسياً وأفقياً فى الخانة (د) . نكرر نفس الطريقة لكل حالة لها قيمتين فنحصل على الجدول الانتشارى السابق والذى يتضح منه أن إتجاه العلامات يبتدىء قطرياً وتتجه ناحية اليسار من أعلى إلى أسفل موازياً للسهم الموضح ويطلق على نوع العلاقة تامة موجبة .

ملاحظات عامة

أوضحنا - فيما سبق - مدى أهمية الجداول الإحصائية بأنواعها كأداة لتمثيل وتبويب وتلخيص البيانات والنتائج التى تسفر عنها البحوث الاجتماعية . وتعتمد بنية الجدول الإحصائى على هدف الباحث والغرض الذى من أجله يتم اختيار نوع الجدول وأيضاً أسلوب تبويب البيانات البحثية . وعندما تتعدد الجداول الإحصائية كما وكيفا ، يصبح من الضرورى أن يراعى الدارس عدداً من الملاحظات الهامة والتى من بينها:

أ - فى حالة إعداد تقرير بحثى يتضمن عدداً كبيراً من الجداول . يجب ترقيم كل جدول على حدة . هذا ويفضل أن يكون ترقيم الجدول بإعطائه رقماً تسلسياً وفق ترتيب الجدول فى الفصل . وفى نفس الوقت يقوم الباحث بوضع رقم كل فصل فى حالة تعدد فصول التقرير البحثى. من ثم يفضل وضع رقمين لكل جدول بحيث يكون الرقم الأول من جهة اليمين دالاً على ترتيب الفصل أو نفس رقم الفصل أما الرقم الثانى والذى يفصله عن الرقم الأول مسافة أفقية صغيرة جداً، فتمثل رقم الجدول ويكون الرقم (٢ - ١) دالاً على أنه الجدول رقم (١) فى الفصل الثانى من التقرير .

ب - ضرورة التقيد بكتابة عنوان واضح ومختصر للجدول الإحصائى بحيث يوضح بإيجاز المتغيرات (متغير أو أكثر) التى تمثلها البيانات الإحصائية الذى يتضمنها الجدول فى كل الصفوف والأعمدة . وفى الحالات التى يكون فيها

العنوان عبارة طويلة ، يقوم الباحث باستكمال تفصيل العنوان أسفل الجدول وبعد نهايته يضع علامة مثل ★ أو النقطة المستديرة فى أعلى الجدول ونظيرتها أسفل الجدول بما يفيد إكمال العنوان . كما قد يلجأ الباحث إلى استخدام نفس العلامات أو مثيلاتها من علامات يختارها للتنويه فى هامش الصفحة التى تشتمل على الجدول عن مصادر البيانات أو مناقشتها وخلافه وفق هدف البحث أو موضوع الفصل .

جـ – الحرص على تحديد نوع الوحدة القياسية للمعلومات والبيانات التى يتضمنها الجدول مثل وحدات الأطوال (سنتيمتر ، متر ، كيلومتر وغيرها) أو وحدات الوزن (كيلوجرام ، رطل ، طن وغيرهم) وكذلك وحدات العمر (سنة أو يوم أو شهر) . كما يراعى فى بعض الأحيان ضرورة تحديد اتجاهات القيم بوضع إشارات جبرية سالبة أو موجبة أمام تلك الأرقام . كذلك فى حالة إستخدام النسب فمن الضرورى تحديد تلك النسب هل هى مئوية أو فى الألف أو فى المليون وهكذا.

د – يجب أن تكون عناوى الصفوف والأعمدة مختصرة وواضحة .

الفصل الثالث

التمثيل البيانى للبيانات

نظـــام المحاور الأحــــداثيـة:

التمثيل البيانى للبيانات المتقطعة :

١- المستطيلات أو الإعمدة.

(أ) الأعمدة البسيطة.

(ب) الأعمدة المجزأة.

(جـ) الأعمدة المزدوجة.

(د) الأعمدة المنزلقة.

٢- الرسوم الدوائرية والقطعية.

التمثيل البيانى للبيانات والتوزيعات التكرارية المتصلة :

المدرج التكرارى.

المضلع التكرارى.

المضلع التكرارى التجمعى.

المنحنى التكرارى.

المنحنيات المتجمعة:

المنحنى المتجمع الهابط.

المنحنى المتجمع الصاعد.

الرسومات البيانية المبعثرة



الفصل الثالث

التمثيل البيانى للبيانات

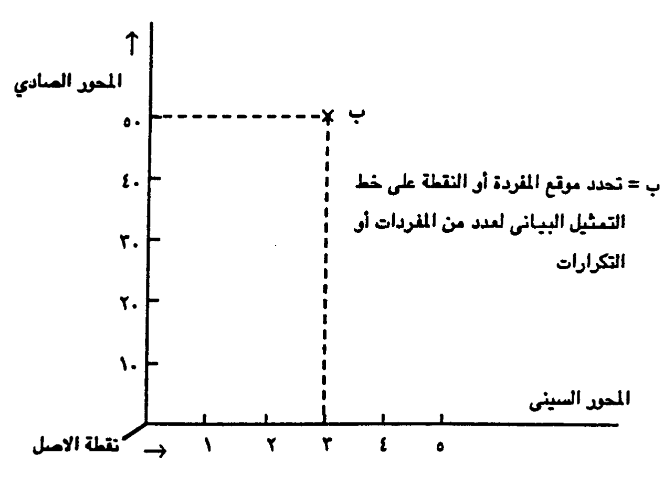
غالباً ما يتردد بعض الباحثين فى قراءة الجداول ويفضلون عليها تمثيل البيانات الجدوليه باستخدام الرسومات البيانية والتى تعطى لهم تطوراً وتفهماً أكبر وأوضح مما تعطيه الجدوال التكراريه. وتعتبر من أبسط طرق التمثيل البيانى تلك التى تقوم على التباينات أو الاختلاف بين قيم التكرارات وذلك باستخدام المساحات أو الارتفاعات النسبية لقمم المنحنيات. مع مقارنة التكرارت بهذه الطريقة لكل مفردة أو تصنيف. فارتفاع العمود أو الخط الرأسى وكذلك المستطيل يحدد حجمه النسبى. فلو كان المقياس نوعياً Nominal، فإن الترتيب الواقعى للأعمدة البيانية لا يكون مألوفاً أو مقبولاً. أما فى حالات المقاييس الترتيبية والكمية يسهل على الباحث التمثيل بأعمدة بيانية يسهل رسمها فى شكل ترتيبى وبطريقة تعطى دلالة ذات معنى عن طبيعة التوزيعات التكرارية.

ولكن قبل أن نتناول الأشكال البيانية المختلفة للتوزيعات التكرارية نرى ضرورة الشرح بإيجاز لنظام المحاور الأحداثية المشتركة، والذى يتضمن مستويين أو محورين أحداهما يسمى المحور الأفقى أو المحور السينى بينما يطلق على المحور الرأسى المحور الصادى.

نظام المحاور الاحداثية : Cartesian Coordinate System :

ويتكون هذا النظام المشترك من مقياسين رقميين يتعامد كل منهما على الآخر بزاوية قدرها ٩٠ درجة، ويطلق على المقياس الأول المقياس السينى. ويتمثل بخط أفقى يتم تقسيمه إلى مسافات متساوية الأبعاد ابتداءاً من نقطة الالتقاء بين هذا المقياس والمقياس الرأسى المتعامد عليه ويطلق على نقطة الالتقاء نقطة الأصل Origin، كذلك يتم تقسيم المحور الرأسى إلى مسافات متساوية الأبعاد ويطلق على المقياس الأفقى المحور السينى وعلى المقياس الرأسى المحور الصادى. ومعنى نظام محاور مشتركة إحداثية أى أن المفردة الواحدة يتم توقيعها وفق قيمتين مناظرتين

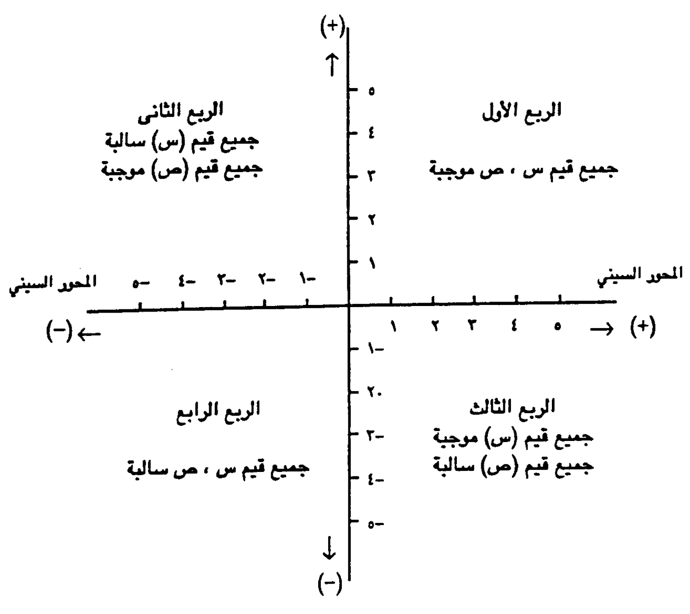
لها حيث أن لكل قيمة على المقياس السينى (قيمة مستقلة) وقيمة أخرى مناظرة تابعة على المحور للصادى الرأسى (قيمة معتمدة) وطبقاً للارتباط بين القيمتين السينية والصادية للمفردات يتحدد شكل المنحنى أو الرسم البيانى ويحصل على علاقة أما خطية أو انحنائية بين المتغيرين (س،ص) . حيث (س) = طول المسافة ابتداء من نقطة الأصل حتى قيمتها المعطاة . وأيضاً (ص) = طول المسافة ابتداء من نقطة الأصل حتى قيمتها الصادية المعطاة كما يتضح ذلك من الشكل رقم (٣-١) بالنسبة لمفردة واحدة تحددها قيمتان (٥٠،٣) وطريقة تحديد موضع المفردة بإحداثيها (٥٠،٣) . أن نبدأ فى قياس ثلاث وحدات على المحور السينى



شكل رقم (٣ - ١)

بداية من نقطة اليسار وعند القيمة (٣) يقام خط رأسى موازى للمحور الصادى. ثم بعد ذلك نقوم بقياس القيمة (ص) وهى (٥٠) لنفس المفردة على المحور الصادى وعند تلك القيمة نرسم خط أفقى موازى للمحور السينى فيلتقى الخطان فى نقطة (ب) كما فى الشكل السابق ، تحدد موضع المفردة المعطاة . وفى حالة تعدد المفرادات أو التكرارت نكرر نفس العمل لكل مفردة فنحصل فى النهاية على الشكل البيانى أو التمثيل البيانى للتوزيعات التكرارية .

وإذا كانت قيمتا المفردة وهما (س،ص) سالبتين أو موجبتين أو تكون أحداهما سالبة والأخرى موجبة، فإن المحاور الإحداثية تقسم إلى أربعة مساحات - اثنتين أعلى المحور الأفقى بحيث تكون الأولى ناحية اليمين من نقطة الأصل جميع قيمها موجبة والثانية إلى اليسار من نقطة الأصل ولها نفس المقياس بأبعاده المتساوية وتكون جميع القيم الواقعة بداخلها سالبة للقيمتين (س،ص) لكل



شكل رقم (٣ - ٢) محاور الاحداثيات فى مستويين (س ، ص)

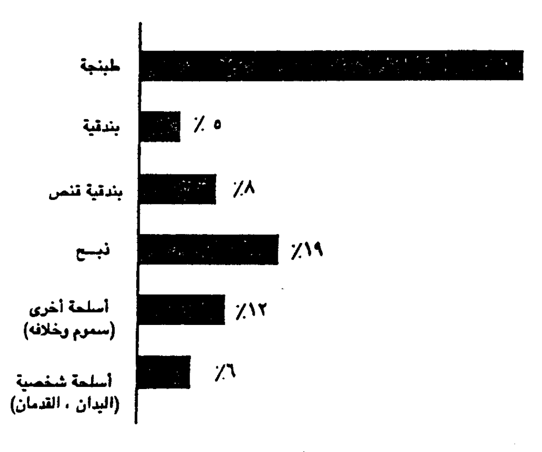
مفردة كذلك توجد مساحتان أسفل الخط الأفقى يقسمهما امتداد المحور الصادى متجهاً إلى أسفل من نقطة الأصل بحيث تكون جميع قيم (س) موجبة وجميع قيم (ص) سالبة للمفردات وهذه المساحة تقع على يمين المحور الصادى المتجه لأسفل. أما المساحة الرابعة فتقع أسفل المحور الأفقى وعلى اليسار من امتداد المحور الصادى المتجه إلى أسفل كما يتضح ذلك من الشكل رقم (٤-٢) . وفى هذه المساحة الرابعة تكون جميع قيم (س) وجميع قيم (ص) سالبة.

التمثيل البيانى للبيانات المتقطعة Discrete data

ذكرنا – فيما سبق أنه يمكن التعبير عن المتغيرات المتقطعة أما بقياسات تصنيفية Nominal أو بقياسات ترتيبية Ordinal ففى حالة القياس التصنيفى، يتم تقسيم المتغير إلى أجزاء نوعية متشابهة تبعاً للصفة تحت الدراسة تميزها مجموعة فرعية عن أخرى وفى حالة استخدام المقياس الترتيبى يكون ترتيب المجموعات وفقاً للتميز فى الصفة بحيث نقول على سبيل المثال أن المجموعة رقم (١) فرضاً أكبر من (أو أقل من) المجموعة رقم (كذا) . أيضاً يمكن استخدام الأشكال البيانية التالية فى تمثيل التوزيعات التكرارية المتقطعة.

المستطيلات أو الأعمدة:

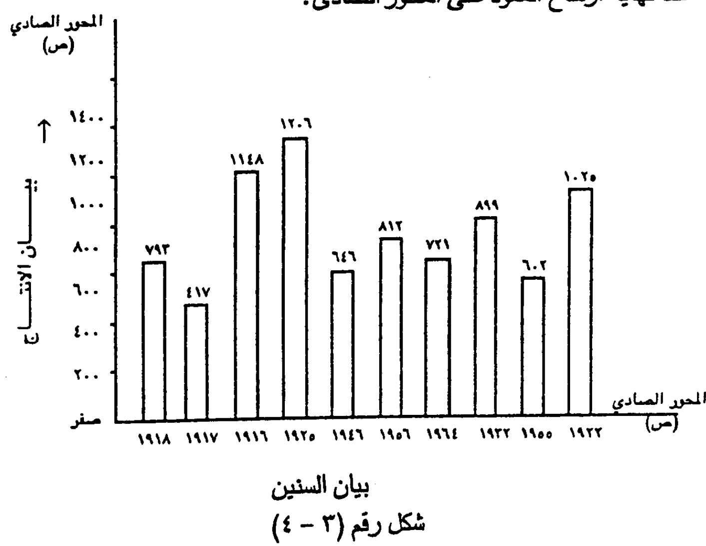
نظراً لعدم معرفة الفروق أو المسافات بين القيم المختلفة للبيانات التصنيفية والترتيبية، نجد أن التمثيل البيانى لها مقيد بنسب الحالات فى شكل ترتيبى أو تصنيفى وفقاً لنوع المقياس المستخدم. وتستخدم المستطيلات أو الأعمدة فى تمثيل النسب المئوية للقيم المختلفة. وعلى سبيل المثال، يوضح الشكل رقم (٣-٣) استخدام الأعمدة فى تمثيل نسب الجرائم تبعاً لنوع السلاح المستخدم بيانياً.



شكل رقم (٣- ٣) التمثيل البيانى بالأعمدة يوضح النسب المئوية لجرائم القتل تبعاً لنوع السلاح المستخدم فى ارتكاب الجريمة

الأعمدة البسيطة :

تتناول تلك الطريقة ظاهرة ذات متغيرين فى أغلب الأحوال مثال ذلك إنتاج القطن المصرى فى خلال عشرة سنوات. حيث تمثل السنوات متغيراً مستقلاً، بينما انتاج القطن يمثل متغيراً تابعاً ، ولكى يتم تمثيل ذلك بيانياً يستخدم ورقة الرسم البيانى ويرسم محورين أحدهما يمثل المحور السينى والثانى يمثل المحور الصادى حيث يتقاطعان فى نقطة تسمى نقطة الأصل ثم نقوم بتقسيم المحور السينى إلى مسافات متساوية تمثل عدد السنوات ويمثل المحور الصادى بمقياس رسم مناسب حجم الانتاج من القطن كما هو موضح بالشكل رقم (٣-٤). وفى حالة وجود رقم شاذ يصعب تدوينه على المحور الصادى مثلاً لصغر حجم ورقة الرسم البيانى يقوم الدارس بعمل تقطيع على العمود بشكل شرشرة متوازية ثم يكتب الرقم ثم يكتب عند نهاية ارتفاع العمود على المحور الصادى.

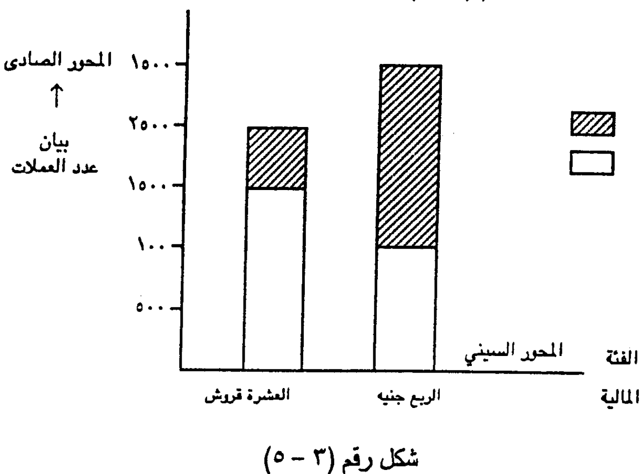


شكل رقم (٣ - ٤)

الأعمدة المجزأة :

وتستخدم فى حالة تمثيل متغير واحد يراد إظهار مكوناته الداخليه أو عناصره الفرعية. فمثلاً عند قيام أحد الصيارفة بجرد محتويات خزينة النقود التى بعهدته. وإن المحتويات تشتمل على عملات معدنية وأخرى ورقية ذات فئات مالية

معروفة. ففى هذه الحالة يمكن استخدام الأعمدة المجزئة حيث تكون النقود إجمالاً هى المتغير ومحتوياتها الفضية والورقية هى العناصر الفرعية المكونة لها. أو بمعنى آخر يمثل المستطيل المتغير ويقسم بنسب المحتويات إلى أجزاء كما يتضح ذلك من الشكل التالى رقم (٣ -٥).



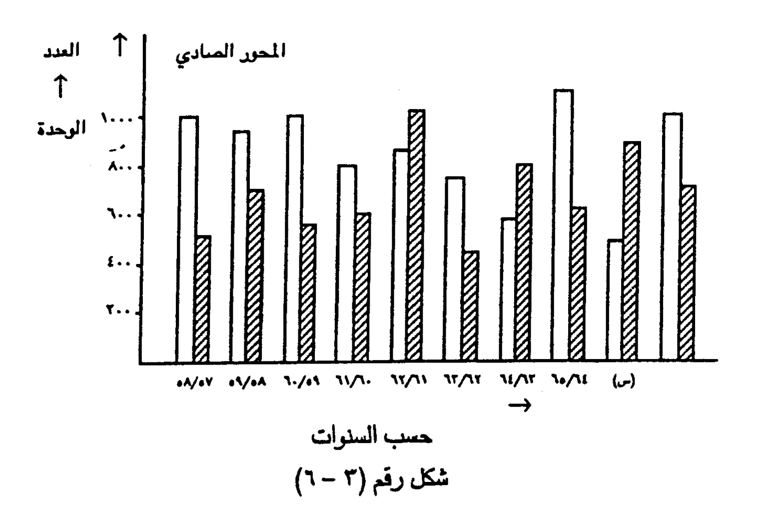
شكل رقم (٣ - ٥)

الأعمدة المزدوجة أو المتلاصقة Coponent Bar Chart :

يستخدم هذا النوع من الأعمدة فى التمثيل البيانى للظاهرة إذا تضمنت متغيراً تابعاً Dependent Variable ينقسم داخلياً إلى قسمين أو اكثر.

مثال (١) :

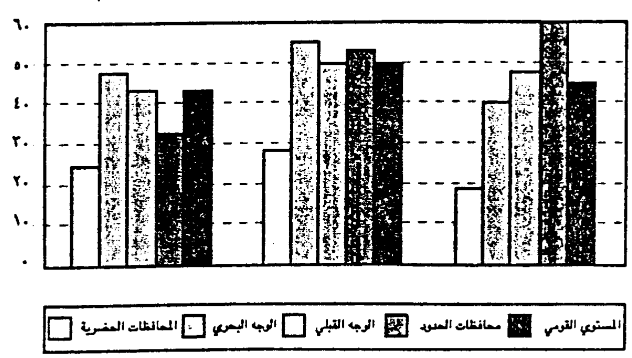
يوضح الشكل البيانى رقم (٣-٦) التمثيل بالأعمدة المزدوجة لنسب عدد المواليد والوفيات لكل عام ولفترة عشر سنوات. ويلاحظ فى الشكل البيانى أن قاعدة المستطيل تنقسم إلى قسمين على النحو المبين بالرسم. أيضاً يوضح الشكل رقم (٣-٧) معدلات التصويت حسب الأقاليم فى جمهورية مصر العربية باستخدام الأعمدة المتلاصقة .



حسب السنوات

شكل رقم (٣ - ٦)

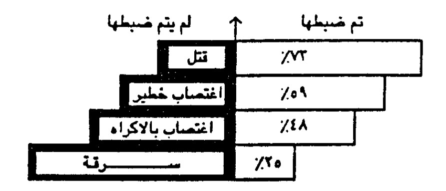
معدلات التصويت حسب الأقاليم (انتخابات مجلس الشعب)



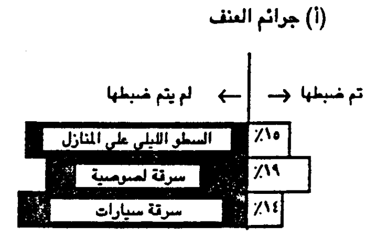
الشكل رقم (٣ - ٧)

الاعمدة المنزلقة Sliding Bar Charts :

وتستخدم فى التمثيل البيانى لمتغيرين أحدهما ينقسم إلى شعبتين أو قسمين. ومن ثم نجد تشابهاً فى الاستخدام بين الأعمدة المنزلقة والأعمدة المجزأة إلا أن طريقة الرسم البيانى تختلف فى حالة الأعمدة المنزلقة . ولرسم الأعمدة المنزلقة يقوم الدارس برسم محور رأسى فى منتصف الورقة تقريباً على أن تكون قيم الشعبة الأولى للمتغير على الجانب الأيمن من هذا المحور كما تكون قيم الشعبة الثانية لنفس المتغير المنقسم على الجانب الأيسر. وفى هذه الطريقة، تمثل أطوال المستطيلات أو الأعمدة الأفقية نسب الحالات الواقعة إلى اليمين واليسار من المحور الصادى كما يتضح ذلك من التمثيل البيانى رقم (٣ - ٨) لنسب الجرائم خلال عام ١٩٧٩ التى تم التحقيق فيها وأيضاً التى لم يتم التوصل فيها إلى الجانى داخل الولايات المتحدة الأمريكية وفق التقرير السنوى للاتحاد الفيدرالى الذى نشر فى أوائل عام ١٩٨٠.



(أ) جرائم العنف

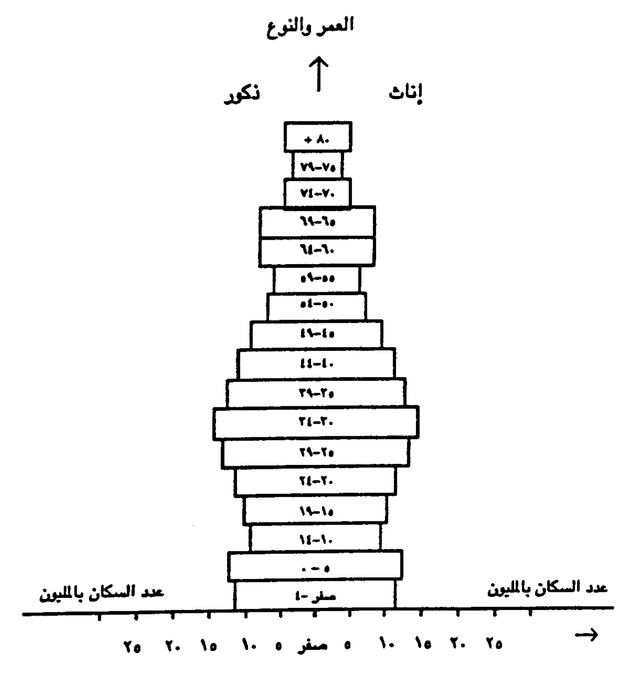


(ب) جرائم ضد الممتلكات

شكل رقم (٣ - ٨) استخدام الأعمدة المنزلقة فى التمثيل البيانى لنسب الجرائم خلال عام ١٩٧٩

ويتضح من التمثيل البيانى مدى التباين الكبير بين المقبوض عليهم فى الأنواع المختلفة من الجرائم الموضحة بالتقرير السنوى. وتعتبر جرائم القتل أكثر أنواع الجرائم وقوعاً فى أيدى البوليس بينما نسب جرائم الاعتداء على الممتلكات هى الأقل ضبطية.

أيضاً يشيع التمثيل البيانى باستخدام الأعمدة المنزلقة فى حالات التوزيعات التكرارية لتصنيفات العمر والنوع Sex فى مجال الإحصاء الحيوى أو السكانى، وتأخذ الأعمدة المنزلقة شكلاً هرمياً، ويطلق على التمثيل البيانى الهرم السكانى Population pyramid نظراً لارتباط الشكل البيانى الهرمى بخاصية المتغير موضوع البحث. ففى شكل رقم (٣ - ٩) يوضح التمثيل الهرمى السكانى لتوزيع السكان ذكوراً وأناثاً وفق الفئة العمرية فى أحدى الدول .

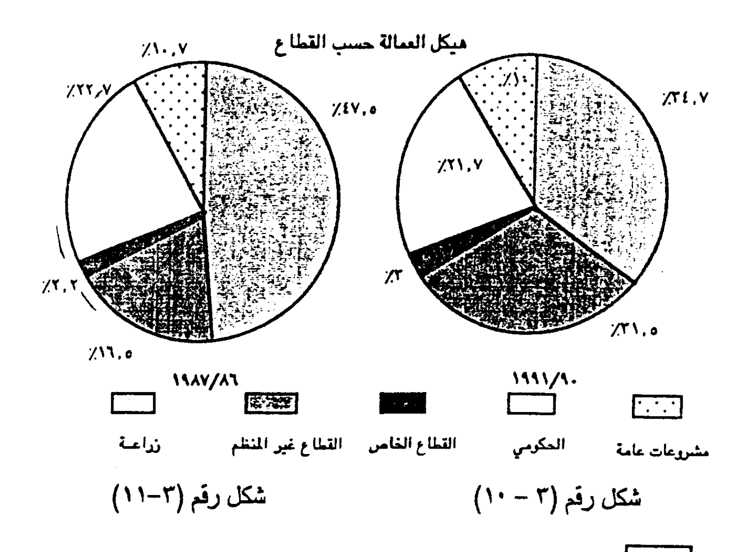


شكل رقم (٣ – ٩) التمثيل الهرمى لتوزيع السكان فى أحدى الدول

الرسوم الدوائرية والقطعية Pie and Circular charts :

تعتبر الدوائر بديلاً آخراً لتمثيل البيانات المتقطعة. كما تعتبر الرسوم الدوائرية أول وأبسط أشكال التمثيل البيانى والتى غالباً ما تستخدم إذا كانت النسبة المئوية للمجموع الكلى للقيم المقاسة تقع داخل كل من تصنيفات المتغير وذلك بتقسيم الدائرة إلى أجزاء وفق تلك النسب. ولما كان استخدام الرسوم الدوائرية للمجموع الكلى المقسم داخلياً إلى تصنيفات، فلا يفضل استخدام تلك الرسوم إذا كان عدد تلك التصنيفات الفرعية كثيراً أو متعدد الأوجة خشية عدم القدرة على التمييز بينها بسهولة. ومن ثم يفضل استخدام الأعمدة بدلاً من الرسوم الدوائرية فى هذه الحالة.

وعندما نقسم الدائرة بخطوط تبدأ جميعها من المركز فى المنتصف فإننا سنحصل على قطع من الدائرة تتباين فى مساحتها وفقاً للنسب المئوية وللزواية المركزية لكل قطعة. فكما نعلم أن الزاوية المركزية للدائرة قدرها ٣٦٠ درجة فإن كل (١ ٪) يقابله جزء من الزاوية المركزية قدره (٣,٦) درجة، كذلك عند استخدام التمثيل الدوائرى يراعى ترتيب النسب الفرعية للمجموع الكلى فى شكل تسلسل أما تصاعدياً أو تنازلياً كما يتضح من الشكلين رقم (٣-١٠) ، (٣-١١) :



شكل رقم (٣-١١)

شكل رقم (٣ - ١٠)

وتوجد طريقة أخرى لرسم الدوائرة وهى عن طريق استخدام نصف قطر الدائرة وسنوضح ذلك من خلال المثال التالى:

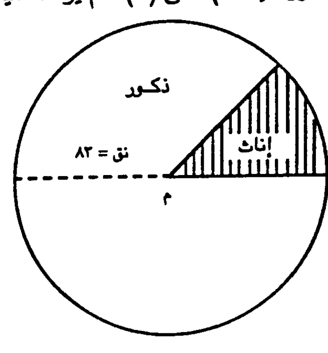
مثال (٢) :

بلغ عدد الذكور والإناث من البالغين فى إحدى المراكز التى تتسم بمعدل عالٍ لجذب السكان (٢٥٩٠٧ ذكراً) ، (٤٣٨٥ أنثى) وذلك خلال عام ١٩٩٦ . والمطلوب استخدام الدائرة فى تمثيل تلك البيانات.

خطوات الحل:

١ – تقسيم الدائرة داخلياً حسب أعداد الذكور والإناث بما يتناسب وإجمالى درجات الدائرة (٣٦٠درجة) .

٢ – حساب نصف قطر الدائرة (نق) وهو عبارة عن خارج قسمة الجذر التربيعى للمجموع الكلى (ذكور+ إناث) على (٢) . ثم يؤخذ مقياس رسم مناسب.



شكل رقم (٣ – ١٢)

٣ – لتقسيم الدائرة إلى أناث وذكور، يلزم أن تنسب قيمة كل شريحة إلى إجمالى الشريحتين مضروبة × ٣٦٠ بهدف تحديد نصيب كل شريحة من الدرجات.

حل المثال:

الذكور	الإناث	المجموع
٢٥٩٠٧	٤٣٨٥	٣٠٢٩٢

$$\text{نق} = \frac{\sqrt{30292}}{2} = 87$$

يؤخذ مقياس رسم مناسب وليكن كل سم يساوى ٢٥

نق = ٣,٤٨

عدد الذكور = $\frac{٢٥٩٠٧}{٣٠٢٩٢} \times ٣٦٠ = ٣١٠$ درجة

عدد الإناث = $\frac{٤٣٨٥}{٣٠٢٩٢} \times ٣٦٠ = ٥٠$ درجة

التمثيل البياني للبيانات والتوزيعات التكرارية المتصلة :

Continuouos Frequency Data

أوضحنا في الفصل الثاني أنه يمكن تلخيص البيانات المتصلة بقياسات كمية بواسطة جداول إحصائية أما تكرارية أو مزدوجة أو متجمعة. ويمكن التعبير أيضاً عن تلك البيانات التكرارية المتصلة باستخدام الأشكال والرسوم البيانية التالية:

المدرج التكرارى Frequency histogram :

يمكن تقسيم التمثيل البيانى بالمدرج التكرارى وفق التكرارات إلى نوعين الأول المدرج التكرارى والثانى هو المدرج التكرارى النسبى. وكلا النوعين من المدرج التكرارى يستخدمان فى البيانات الكمية. والخطوة الأولى لرسم المدرج التكرارى تتمثل فى ترتيب وتنظيم البيانات المعطاة فمثلاً لو افترضنا أن لدينا ١٠٠ طفلاً مولودا فى إحدى المستشفيات ويتم باستمرار متابعة الزيادة فى أوزانهم شهرياً وفق قائمة معينة من الغذاء وإن المتابعة تستمر حتى الشهر السادس وأمكن الحصول على بيانات الوزن الزيادة لهؤلاء الأطفال (بالكيلو جرام) .

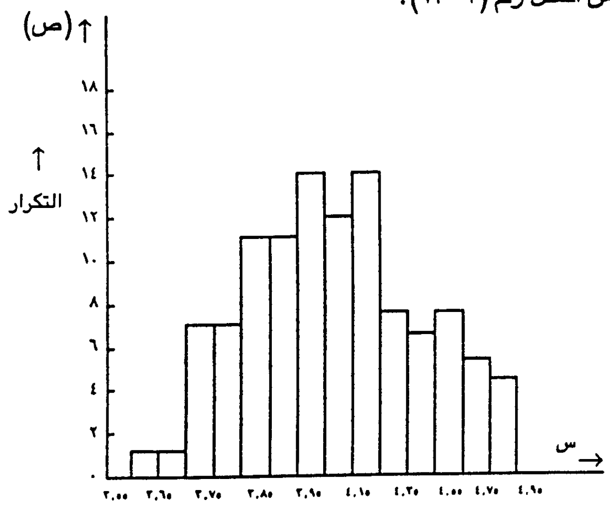
٣,٧	٤,٢	٤,٤	٤,٤	٤,٣	٤,٢	٤,٤	٤,٨	٤,٩	٤,٤
٤,٢	٣,٨	٤,٢	٤,٤	٤,٦	٣,٩	٤,٣	٤,٥	٤,٨	٣,٩
٤,٧	٤,٢	٤,٢	٤,٨	٤,٥	٣,٦	٤,١	٤,٣	٣,٩	٤,٢
٤,٠	٤,٢	٤,٠	٤,٥	٤,٤	٤,١	٤,٠	٤,٠	٣,٨	٤,٦
٤,٩	٣,٨	٤,٣	٤,٣	٣,٩	٣,٨	٤,٧	٣,٩	٤,٠	٤,٢
٤,٣	٤,٧	٤,١	٤,٠	٤,٦	٤,٤	٤,٦	٤,٤	٤,٩	٤,٤
٤,٠	٣,٩	٤,٥	٤,٣	٣,٨	٤,١	٤,٣	٤,٢	٤,٥	٤,٤
٤,٢	٤,٧	٣,٨	٤,٥	٤,٠	٤,٢	٤,١	٤,٠	٤,٧	٤,١
٤,٧	٤,١	٤,٨	٤,١	٤,٣	٤,٧	٤,٢	٤,١	٤,٤	٤,٨
٤,١	٤,٩	٤,٣	٤,٤	٤,٤	٤,٣	٤,٦	٤,٥	٤,٦	٤,٠

فى هذه الحالة لابد من عمل جدول تكرارات حتى يسهل عمل المدرج التكرارى . وبنفس خطوات عمل الجدول التكرارى السابق شرحه فى الفصل الثانى - تحصل على جدول التكرارات رقم (٣ - ١)

جدول رقم (٣ - ١)

الفئة	طول الفئة	التكرارات	التكرارات النسبية
١	٣,٥٥ -	١	٠,٠١
٢	٣,٦٥ -	١	٠,٠١
٣	٣,٧٥ -	٦	٠,٠٦
٤	٣,٨٥ -	٦	٠,٠٦
٥	٣,٩٥ -	١٠	٠,١٠
٦	٤,٠٥ -	١٠	٠,١٠
٧	٤,١٥ -	١٣	٠,١٣
٨	٤,٢٥ -	١١	٠,١١
٩	٤,٣٥ -	١٣	٠,١٣
١٠	٤,٤٥ -	٧	٠,٠٧
١١	٤,٥٥ -	٦	٠,٠٦
١٢	٤,٦٥ -	٧	٠,٠٧
١٣	٤,٧٥ -	٥	٠,٠٥
١٤	٤,٨٥ -	٤	٠,٠٤
المجموع		١٠٠	١,٠٠

بعد ذلك يتم استخدام محاور الاحداثيات (س، ص) ونختار المحور السينى للفئـات والمحـور الصـادى للتكرارت. ويتكون المدرج التكرارى بواسطة رسم مجموعة من المستطيلات تكون طول قاعدتها على المحور السينى مساوياً لطول الفئة وفى هذا المثال نجد أن طول الفئة ثابت ومقداره ١٠، ويأخذ مقياس رسم مناسب حيث أن طول الفئة صغير ففى هذه الحالة يمكن ضرب طول الفئة × ١٠ حتى يسهل توقيع القيمة على المحور وينوه الدارس عن ذلك بأن طول الفئة مضروباً فى ١٠ . أما ارتفاع المستطيل فقيمته هى التكرار المناظر لكل فئة ومن ثم نجد أن مساحة كل مستطيل تتناسب طردياً مع التكرار المناظر له. وبالتالى نحصل على عدة مستطيلات متجاورة ويطلق على هذا الشكل المدرج التكرارى كما يتضح ذلك من الشكل رقم (٣- ١٣):

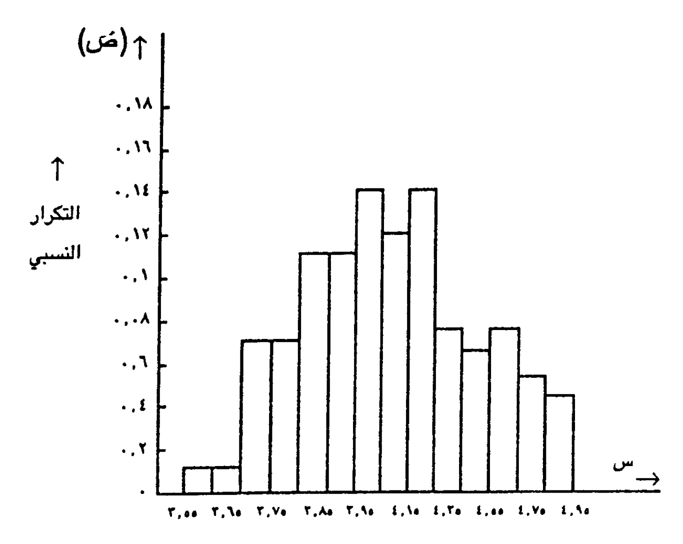


→ الفئات (الزيادة فى وزن الطفل)

شكل رقم (٣ – ١٣) المدرج التكرارى لبيانات أوزان الاطفال

كذلك يمكن رسم المدرج التكرارى النسبى. وكما نعرف أن التكرار النسبى لأى فئة بأنه تكرار هذه الفئة على المجموع العددى الكلى للتكرارات على طول المدى ، ومن ثم اضفنا عمود للتكرارات النسبية فى جدول التكرارات السابق

(العمود الأخير) . ويمكن أيضاً بنفس الخطوات رسم المدرج التكراري النسبي في المثال السابق كما يوضح ذلك الشكل رقم (٣-١٤) .



← الفئات (الزيادة في وزن الطفل)

شكل رقم (٣ - ١٤) المدرج التكراري النسبي لبيانات أوزان الأطفال

خطوات رسم المدرج التكراري :

١- تستخدم محاور الإحداثيات (س،ص) بحيث يستخدم المحور الأفقي (السيني) في تمثيل الفئات وذلك بتقسيمه إلى أقسام متساوية بمقياس رسم مناسب. كما يستخدم المحور الصادي في تمثيل التكرارت أو التكرارات النسبية ويقسم أيضاً إلى أقسام متساوية حتى أكبر قيمة للتكرارات أو أكبر نسبة تكرارية. ولايشترط أن نستخدم نفس التقسيم للمحورين . المهم أن يستوعب كل محور كل الفئات بالنسبة للسيني وكل التكرارات حتى أكبرها بالنسبة للمحور الصادي كما يراعى أيضاً أن نستخدم تقسيماً واحداً نختاره أو مقياس رسم واحد لكل محور منهما.

٢ – يتم تمثيل التكرارات باستخدام أعمدة بحيث أن الضلعين الرأسيين لقيم التكرارت يبدأ أولهما من الحد الأدنى للفئة والثانى من قيمة الحد الأعلى للفئة ذلك على المحور الأفقى أما الارتفاع لهما فيساوى قيمة التكرارات لنفس الفئة.

٣ – إذا كانت الفئات متساوية الطول فى التوزيع التكرارى فإن النسبة بين ارتفاعات المستطيلات تعادل النسب بين مساحاتها بحيث أن المجموع المساحى لكل المستطيلات يعتبر مجموع التكرارات الكلى.

٤ – إذا كانت الفئات غير متساوية الطول فيلزم ضرورة إجراء تعديل للتكرارات وذلك بقسمة تكرار كل فئة على طولها فنحصل على تكرار معدل يمثل فى هذه الحالة ارتفاع المستطيل أما قاعدته فتختلف فى قيمتها من مستطيل لآخر نظراً لاختلاف طول الفئات.

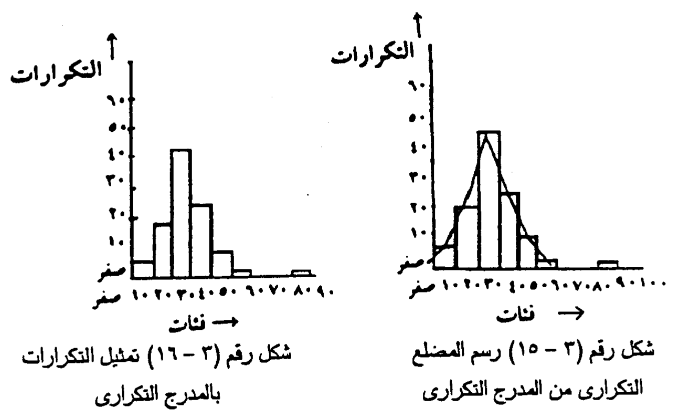
٢- المضلع التكرارى Frequency polygon :

ويمثل البديل الثانى لتمثيل التوزيعات التكرارية المتصلة بيانياً حيث يشابه إلى حد بعيد المدرج التكرارى كما قد يستنبط منه. وبدلاً من أن يعتمد المدرج التكرارى على توزيعات تكرارية مقفلة حيث يتحدد حدى الفئة (الأدنى والأعلى) ويمثلان ضلعى المسستطيل، فإن رسم المضلع التكرارى يقوم على فكرة استخدام مركز الفئة أو القيمة المتوسطة لها. ومن ثم يعتمد المضلع التكرارى على نقطة فى مركز الفئة وهذه النقطة أما تتحدد فى المدرج التكرارى بتنصيف للخط العلوى لكل مستطيل والموازى للمحور الأفقى فى نقطة تمثل مركز الفئة لكل تكرار مقابل على المحور الصادى. أو يمكن حسابها بقسمة مجموعة قيمتى حدى الفئة (الأدنى + الأعلى) وقسمة الناتج على (٢) وذلك فى حالة عدم وجود المدرج التكرارى. وبتوصيل النقط المتوسطة بخطوط مستقيمة بين كل نقطتين متتاليتين نحصل على شكل المضلع التكرارى.

(أ) رسم المضلع التكرارى من المدرج التكرارى:

الشكل رقم (٣-١٦) يبين المدرج التكرارى لتوزيع تكرارت متصلة ذات أطوال فئات متساوية.

الشكل رقم (٣-١٥) يبين المضلع التكرارى الناتج من تنصيف الضلع الأفقى والعلوى لكل مستطيل والتوصيل بخطوط مستقيمة بين كل نقطتين متتاليتين منهما لنفس التوزيعات التكرارية .



شكل رقم (٣ - ١٥) رسم المضلع التكرارى من المدرج التكرارى

شكل رقم (٣ - ١٦) تمثيل التكرارات بالمدرج التكرارى

(ب) رسم المضلع التكرارى من التوزيعات التكرارية مباشرة :

مثال(٣) :

ارسم المدرج التكرارى للتوزيعات التكرارية لأوزان عدد من الأفراد.

ف	ك
١٩٠ - ٢٠٠	٣
١٨٠ -	٤
١٧٠ -	٥
١٦٠ -	١٥
١٥٠ -	١٨
١٤٠ -	١١
١٣٠ -	٨
١٢٠ -	٤
١١٠ -	٥
١٠٠ -	٢
المجموع	٧٥

الحـــل :

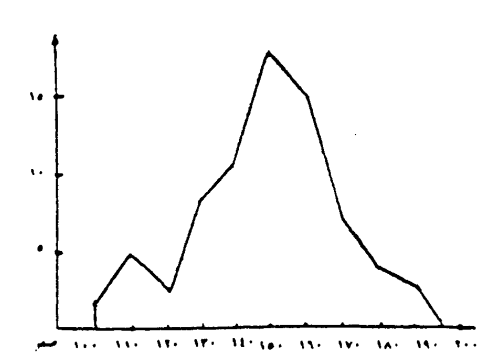
تحسب نقطة المنتصف من حاصل جمع الحد الأعلى + الحد الأدنى لكل فئة وبقيمة هذا المجموع على (٢)

ونكرر هذا العمل لكل فئة . فبالنسبة للفئة الأولى كمثال :

قيمة النقطة المتوسطة = $\frac{١٩٠ + ١٩٩}{٢}$ = $\frac{٣٨٩}{٢}$

= ١٩٤,٥

ونكرر ذلك فنحصل على المضلع التكرارى الموضح بالشكل رقم (٣ – ١٧) التالى :



المضلع التكرارى المقفل لتمثيل البيانات التكرارية مباشرة

شكل رقم (٣ – ١٧)

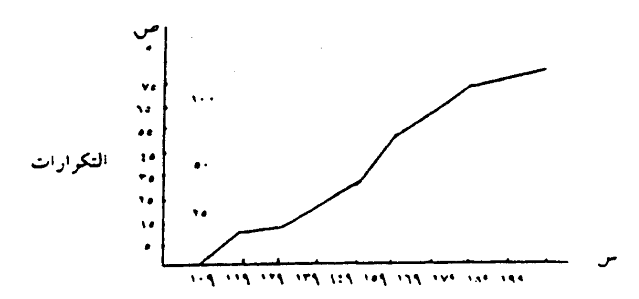
هذا ويفضل دائماً إقفال المضلع التكرارى بمعنى أن نهايتيه تلتقيان مع المحور السينى ويمكن عمل ذلك بأن يفترض الدارس وجود فئتين إضافيتين على الفئات الأصلية التى يشتمل عليها المدى ، شريطة أن يضع أول فئة منهما سابقة لأول فئة أصلية والفئة الثانية يضعها بعد آخر فئة فى المدى. هذا بالإضافة أن يفترض الباحث قيمة تساوى صفراً لتكرارات كل فئة منهما ويشرط أن تتساويا فى

طول الفئة مع باقى الفئات الأصلية. والشكل رقم (٣-١٧) فى المثال السابق يبين المضلع التكرارى المقفل.

بعض استخدامات المضلع التكرارى :

المضلع التكرارى التجمعى Cumulative Frequency Polygon :

يمكن ترتيب التوزيعات التكرارية تجمعياً تصاعدياً أو تنازلياً، ومنها يمكن رسم المضلع التكرارى التجمعى. ويعد هذا الشكل إحدى استخدامات المضلع التكرارى، ففى المثال السابق للتوزيعات التكرارية للأوزان (بالرطل) يمكن التمثيل البيانى بمضلع تكرارى تجمعى صاعد (من أقل قيمة إلى أعلى قيمة). إلا أن طريقة رسم المضلع التكرارى التجمعى لا تعتمد على قيم النقط المتوسطة بل تعتمد على قيم الحد الأعلى لكل فئة وبتوصيل تلك القيم بعضها ببعض باستخدام خطوط مستقيمة فإننا نحصل على المضلع التكرارى التجمعى الموضح بالشكل (٣ - ١٨).

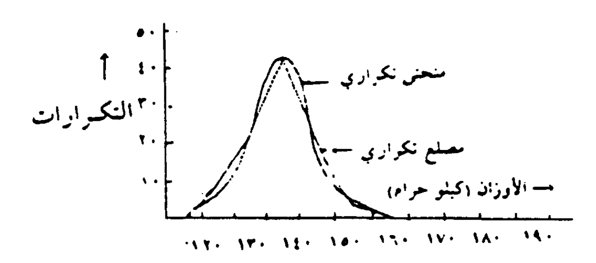


شكل رقم (٣-١٨) المضلع التكرارى التجمعى للأوزان

٢- المنحنى التكرارى Frequency Curve :

فى بعض الأحيان يرغب الباحثون فى التخلص من التعرجات أو الانكسارات التى يتصف بها المضلع التكرارى والتوصل إلى شكل أكثر تمليساً Smoothing وتحويل المضلع إلى منحنى تكرارى. ومن ثم يمكن القول أن المنحنى التكرارى لا

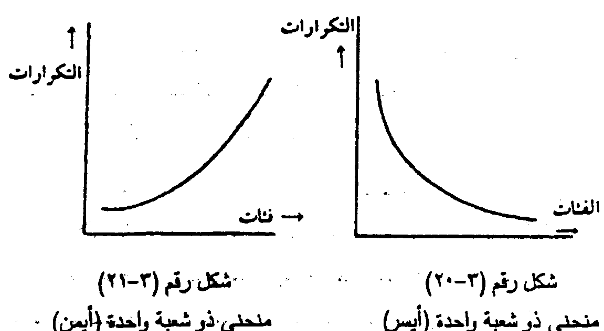
يختلف عن المضلع التكرارى من حيث الشكل إلا فى درجة التمليس كما لا يختلف فى طريقة الرسم. فإذا كنا نقوم بتوصيل النقط المتتالية بخطوط مستقيمة تصل بينها فى المضلع التكرارى ، ففى حالة المنحنى، نقوم باستخدام اليد بتوصيل النقط القريبة من بعضها فى القيم دون الاهتمام بالنقط القريبة منها والشاذة فى قيمتها أحياناً ، سواء كانت تلك النقط الشاذة تعلو أو تقل عن منحنى التوصيل بين النقط القريبة. ومن ذلك لا يمكن القول أن المساحة تحت المنحنى التكرارى تساوى المساحة تحت المضلع التكرارى بل ستكون المساحة الأولى أقل من الثانية كما يتضح ذلك من الشكل رقم (٣-١٩) .



شكل رقم (٣-١٩) رسم المنحنى التكرارى
من المضلع التكرارى للأوزان

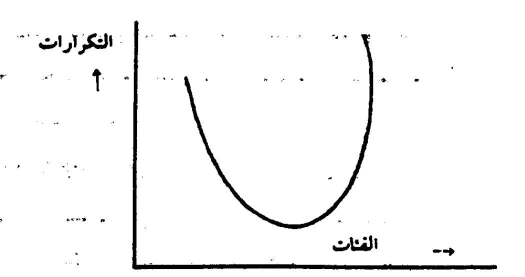
ومن خصائص المنحنيات التكرارية - كما تناولناها سابقاً بالشرح - خاصيتى الالتواء والتفلطح وأيضاً المنحنى الاعتدالى. وفى بعض الحالات الاجتماعية التى تقل فيها عدد التكرارات ، بينما تزيد أعداد أخرى بشكل لا يقارن فإن المنحنى التكرارى يكون فى هذه الحالة ذا شعبة واحدة ، مثال ذلك توزيع السكان على أساس الثروة ، حيث نجد الغالبية العظمى من الفقراء ، بينما قلة قليلة جداً من الأغنياء . بالمثل توزيع أراضى الإصلاح الزراعى فالحيازة قليلة وعدد الحائزين كبير فى حين تقل أعداد الملاك الزارعين أصحاب الأراضى الزراعية كبيرة المساحة . وأمثلة أخرى مشابهة نجد أن المنحنى التكرارى ذا الشعبة الواحدة والموضح فى الشكلين رقم (٣-٢٠) ، (٣-٢١).

فى حالة تجمع التكرارت الكبيرة نسبياً عند طرفى المنحنى كما هو الحال فى تعداد الوفيات على مستوى الجمهورية ، حيث تكثر نسبة الوفيات فى مرحلتى الشيخوخة وسنوات العمر الأولى خاصة السنتين الأولى والثانية من عمر الأطفال، بينما يقل عدد الوفيات نسبياً وبدرجة كبيرة بين الفئة الشبابية ، وفى هذه الحالة نجد أن المنحنى من النوع الناقوسى المقلوب وليس شرطاً أن يكون اعتدالياً ويطلق عليه المنحنى التكرارى ذا الشعبتين كما يوضحه الشكل رقم (٣-٢٢) ويمكن ملاحظة أن تكرارت الفئات الوسطى تقل كثيراً عن تكرارت فئات الطرفين.



شكل رقم (٣-٢٠)
منحنى ذو شعبة واحدة (أيسر)

شكل رقم (٣-٢١)
منحنى ذو شعبة واحدة (أيمن)



شكل رقم (٣- ٢٢) منحنى ذو الشعبتين

المنحنيات المتجمعة Cumulative Frequency :

يعتبر التمثيل البيانى الرابع للتوزيعات التكرارية ويقتصر استخدامه فى حالات تجميع التكرارات فى فئات متتالية. والمنحنيات المتجمعة نوعان إما هابطة أو صاعدة. وتستخدم المنحنيات المتجمعة بنوعيها – كما ذكرنا سابقاً عند شرح الجداول التكراريةالمتجمعة– لنفس الغرض وهو إذا أردنا معرفة عدد أو نسبة المفردات التى تزيد أو تقل عن قيمة أو نسبة معينة ، أو إذا أردنا معرفة الوضع النسبى لقيمة معينة من قيم المتغير.

خطوات رسم المنحنى المتجمع بنوعية الصاعد والهابط:

١ – استخدم المحاور الاحداثية(س،ص) فى الشكل البيانى بأن يكون المحور السينى مقسماً للفئات تقسمات متساوية الأبعاد كذلك تخصيص المحور الصادى للتكرارات أو التكرارت النسبية.

٢ – قم بتوصيل النقط التى تمثل التكرارت المتجمعة عند الحدود العليا للفئات فتحصل على خط ممهد يمثل المنحنى المتجمع الصاعد ولتحقيق ذلك لابد أن تقوم أولاً بتحويل جدول التكرارت المعطى لك إلى جدول تكرارى صاعد. ولما كانت التكرارت المتجمعة عند أى فئة تمثل التكرارت التى تزيد عن الحد الأدنى لتلك الفئة ، فإذا قمت بتوصيل النقط التى تمثل التكرارت المتجمعة عند الحدود الدنيا للفئات فإنك ستحصل على المنحنى المتجمع الهابط. ولتحقيق ذلك أيضاً يجب على الدارس تحويل جدول التوزيع التكرارى المعطى له إلى جدول متجمع هابط بنفس الخطوات السابق شرحها فى عمل جدول التجمع الهابط.

٣ – لا يشترط التقيد بنفس مقياس الرسم للمحورين السينى والصادى ، كما لا يشترط أن يبدأ تقسيم المحور السينى من نقطة الأصل بالقيمة صفر، لتفادى كبر حجم التمثيل البيانى. أما فى حالة رسم المنحنيين الصاعد والهابط معاً على المحاور الإحداثية فيفضل أن تستخدم لهما نفس مقياس الرسم حيث يلتقيان معاً فى نقطة يساوى إحداثيها الصادى نصف مجموع التكرارت.

٤- فى حالة عدم تساوى طول الفئات ، لا يقوم الدارس بإجراء أى تعديل للتكرارات بل يلاحظ فقط صحة رصد القيم التكرارية لحدى الفئة الأدنى والأعلى. والسبب فى عدم الحاجة لتعديل التكرارت أن التمثيل بمنحنى متجمع لا يعدو كونه عملية تجميع فقط.

مثال (٤) :

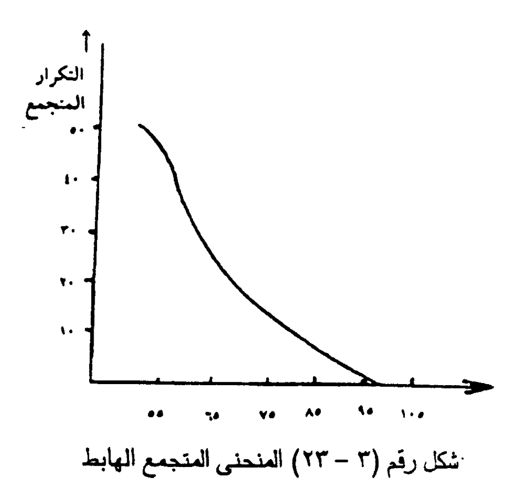
ارسم المنحنين المتجمعين الصاعد والهابط من الجدول التكرارى التالى لدرجات ٥٠ طالباً فى مادة الإحصاء الاجتماعى.

فئات (الدرجات)	التكرار
٥٥ -	٢٢
٦٥ -	١٢
٧٥ -	٧
٨٥ -	٤
٩٥ - ١٠٥	٥
المجموع	٥٠

الحـل :

جدول متجمع هابط

فئات	ك	الحدود الدنيا للفئات	التكرار المتجمع الهابط
٥٥ -	٢٢	٥٥ فأكثر	٥٠
٦٥ -	١٢	٦٥ فأكثر	٢٨
٧٥ -	٧	٧٥ فأكثر	١٦
٨٥ -	٤	٨٥ فأكثر	٩
٩٥ - ١٠٥	٥	٩٥ فأكثر	٥
مجـ	٥٠		

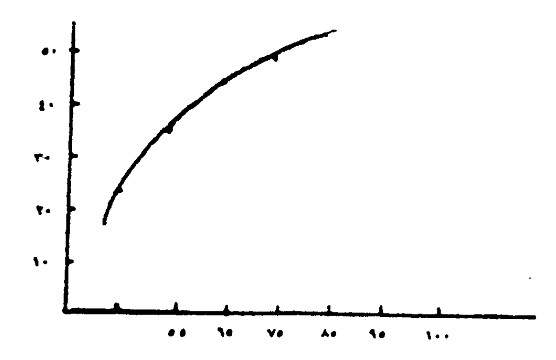


شكل رقم (٣ – ٢٣) المنحنى المتجمع الهابط

جدول متجمع صاعد

فئات	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
٥٥ –	٢٢	أقل من ٦٥	٢٢
٦٥ –	١٢	أقل من ٧٥	٣٤
٧٥ –	٧	أقل من ٨٥	٤١
٨٥ –	٤	أقل من ٩٥	٤٥
٩٥ – ١٠٥	٥	أقل من ١٠٥	٥٠
مجـ	٥٠		

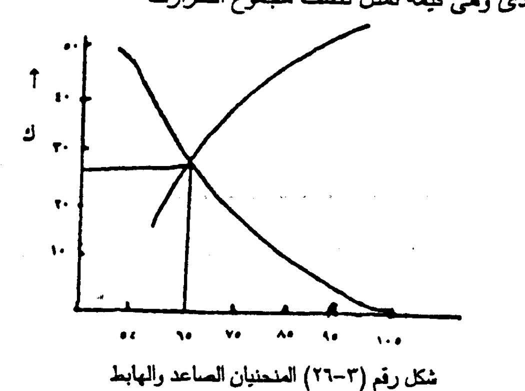
لاحظ أن تكرار الفئة الأخيرة للتكرار المتجمع الصاعد يساوى فى القيمة مجموع التكرارت الأصلية.



شكل رقم (٣-٢٥) المنحنى المتجمع الصاعد

ويمكن رسم المنحنيين فى شكل واحد كما يتضح من الشكل رقم (٣-٢٦):

لاحظ أن نقطة تقاطع المنحنين الصاعد والهابط عند تكرار (٢٥) على المحور الصادى وهى قيمة تمثل نصف مجموع التكرارت



شكل رقم (٣-٢٦) المنحنيان الصاعد والهابط

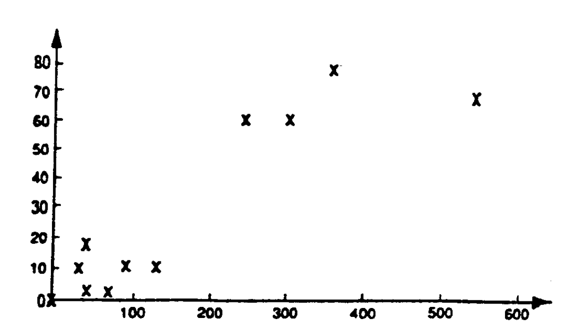
الرسومات البيانية المبعثرة Scattergraphs :

يستخدم هذا النوع من الاشكال البيانية فى تمثيل البيانات المزدوجة Paired data بطريقة يسهل معها على الباحث فحص العلاقات بين زوجين من البيانات والعلاقة بينهما. ويعتبر هذا النوع من الاشكال البيانية التى تعرض القيم المبعثرة خطوة أولى على طريقة فهم ودراسة العلاقات الارتباطية بين متغيرين والانحدار لهما، كما سوف يتضح فى الفصل السادس .

يوضح الشكل رقم (٣-٢٧) أحد الرسومات البيانية المبعثرة لتوضيح العلاقه بين حجم السكان المتوقع عام ٢٠٠٠(بالمليون) والمساحة (لكل ١٠٠كم٢) داخل اثنتى عشرة دولة أوربية كما يتضمنها الجدول الآتى:

الدولة	المساحة (١٠٠٠ كم٢)	حجم السكان المتوقع عام ٢٠٠٠ ميلادية (بالمليون)
بلجيكا	٣٠,٥	٩,٧
الدنمارك	٤٣,١	٥,٢
فرنسا	٥٤٤,٠	٦١,٠
المانيا	٣٥٧,٠	٧٧,٦٥
اليونان	١٣٢,٠	١٠,٠١
ايرلندا	٦٨,٩	٣,٥
ايطاليا	٣٠١,٠	٥٧,٦
لوكسمبرج	٢,٦	٠,٤
هولندا	٤١,٢	١٥,٧
البرتغال	٩٢,١	١١,١
أسبانيا	٥٠٤,٨	٣٨,٧
المملكة المتحدة	٢٤٤,١	٥٨,٨

من بيانات هذا الجدول يمكن رسم الشكل البيانى المبعثر للعلاقه بين السكان والمساحة داخل الدول الاثنى عشر.



شكل (٣ - ٢٧) العلاقة بين حجم السكان والمساحة داخل الدول المختارة

نستخلص من هذا الرسم البيانى الموقع عليه بيانات مبعثرة أن العلاقه طرديه فى اتجاهها بين المساحة وحجم السكان. فالمساحة الكبيرة للدولة يناظرها عدد سكان أكبر ومتوقع لها بحلول عام ٢٠٠٠.

المفاهيم الأساسية Key Concepts

١ - **المتغير** Variable :

هو ظاهرة أو صفة تختلف قيمتها باختلاف الحالات .

٢ - **المتغير المتصل** Continuous Variable :

هو المتغير الذى يأخذ أى قيمة متدرجة على المقياس المستخدم .

المتغير المتقطع Discerte Variable :

هو المتغير الذى يأخذ قيماً محددة أو هو الذى يحتوى مداه على عدد محدود أو لا نهائى من القيم بشرط أن يكون لكل منها قيمة محددة يمكن ترتيبها .

٣ - **المقياس التصنيفى (الاسمى)** Nominal Scale :

هو عملية تصنيف الموضوعات المختلفة إلى فئات تعتمد على سمات محددة .

٤ - **المقياس الترتيبى** Ordinal Scale :

ويتميز عن المقياس التصنيفى بأنه يحتوى على ترتيب منطقى للفئات فضلاً عن اكتسابه صفات هذا المقياس .

المقياس الفئوى Interval Scale :

وهو مقياس يتضمن خصائص المقياسين السابقين ، هذا بالإضافة إلى أن الفروق بين الفئات المختلفة متساوية مع تواجد وحدة القياس ويمكن استخدام العمليات الحسابية فى تحليل البيانات .

التوزيع التكرارى Frequency Distribution :

هو عملية ترتيب الأرقام فى صور تعطى عدد مرات تكرار الرقم فى المجموعة .

جداول التوزيع التكرارى النسبى

: Percentage Frequency Table

يقصد بالتكرار النسبى لفئة ما هو تكرارها بالنسبة إلى التكرار الكلى لجميع الفئات .

جداول التكرار التجمعى Cumulative Frequency Tables :

يستخدم هذا النوع من الجداول التكرارية إذا أراد باحث أن يعرف عدد المفردات التى تقل أو تزيد عن قيمة معينة .

الجداول المزدوجة :

تستخدم فى تلخيص إزدواج القيم لمتغيرين حيث يتم تبويب البيانات وفقاً لفئتين فى ترتيب صفوف وأعمدة بحيث تشمل الصفوف تكرارات الصفة الأولى بينما تشمل الأعمدة تكرارات المتغير الثانى .

تمارين الفصل الثاني والثالث

١- باستخدام الرسوم الدوائرية، وضح النسبة المئوية للانفاق على مجالات الرعاية الصحية الموضحة، والتى تستنفذ ميزانية وزارة الصحة وذلك من الجدول الإحصائى التالى:

مجال الانفاق	النسبة المئوية
رعاية علاجية بالمستشفيات	٣١,٤٠
خدمات طبية	٢٧,٦٩
خدمات أسنان	١٠,٨٤
خدمات مهنية طبية أخرى	٢,٨٤
أدوية ومستلزماتها	١٢,٩٥
نظارات وعدسات طبية	٢,٢٣
رعاية تمريض بالمنزل	٦,٠٩
مجالات أخرى متعلقة بالرعاية الصحية	٥,٩٦

٢- أسفرت نتائج بحث اجتماعى على تبادل الزيارات بين الأصدقاء والأقرباء جاءت استجابات ٨١ مبحوثاً توضح أن الزيارات لا تقل عن مرة واحدة شهرياً والبيانات التالية توضح العدد الفعلى للأشخاص الذين يتم زيارتهم؟

٣	٥	٢	٣	٣	٤	١	٨	٤
٢	٤	٢	٥	٣	٣	٣	١٩	٣
٥	٦	٤	٣	٢	٢	٦	٣	٥
٤	١٤	٣	٤	٦	٣	٤	٢	٤
٩	٤	١	٤	٢	٤	٣	٥	١٥
٤	٣	٥	٧	٣	٥	٦	٢	٢
٥	٤	٢	٣	٦	١	٣	١٦	٥
٣	١١	٤	٥	١٩	٤	٥	٢	٢
٤	٣	١٤	٥	٢	١	٤	٣	٤

المطلوب :

١- عمل توزيع تكرارى من البيانات السابقة.

٢ - إشرح بوضوح اختيارك للفئات.

٣ - بدءاً من إعلان سياسة الانفتاح عام ١٩٧٦ ، أخذت موجات هجرة المصريين خاصة للدول النفطية العربية فى التزايد. ولقد أسفرَت إحدى الدراسات التى أجريت على إحدى مراكز الوجه البحرى لمعرفة النسب المئوية للمهاجرين خلال الفترة من عام ١٩٧٦ حتى عام ١٩٨١ عن النتائج التالية:

السنة	١٩٧٦	١٩٧٧	١٩٧٨	١٩٧٩	١٩٨٠	١٩٨١
النسبة المئوية للمهاجرين بالنسبة للسكان بالمركز	٣,٤	٥,٥	٦,١	٣,٤	٨,٨	١٢,٥

والمطلوب:

١ - هل استخدام الرسم الدوائرى أفضل فى تمثيل البيانات لنسب المهاجرين.

٢ - استخدم المستطيلات فى التمثيل البيانى للبيانات السابقة.

٤ - الجدول التكرارى التالى يتضمن درجات (١٨٠ طالباً) يمثلون درجات طلاب الفرقة الثانية قسم الاجتماع وذلك فى مادة النصوص:

فئات	التكرارات
٨٥-٩٠	٦
٨٠ -	١٥
٧٥ -	٣٧
٧٠ -	٣٠
٦٥ -	٤٢
٦٠ -	٢٢
٥٥ -	١٨
٥٠ -	٧
٤٥ -	٢
٤٠ -	١
المجموع	١٨٠

المطلوب:

- رسم المضلع التكرارى.
- رسم المنحنى التكرارى المتجمع.

٥- البيانات التالية توضح الدرجات التى حصل عليها ثلاثون طالباً فى مادتى الإحصاء والاجتماع الصناعى للفرقة الثالثة قسم الاجتماع:

الإحصاء	الإجتماع الصناعى	الإحصاء	الاجتماع الصناعى	الاحصاء	الاجتماع الصناعى
٦٥	٧٢	٨٥	٨١	٥٩	٧٤
٦٦	٨٦	٧٨	٧٢	٦٢	٨٤
٨٩	٧٥	٧٩	٧٣	٦٨	٨٣
٧٦	٧٧	٧٧	٧٢	٩٦	٨٤
٩٣	٥٩	٧٤	٦٨	٦٤	٦٥
٨٠	٩٢	٨٨	٧٠	٧٧	٧٢
٨٢	٨٠	٦٢	٧٢	٧٢	٨١
٧٤	٨٥	٦٩	٩٥	٥٢	٩١
٨٥	٥٠	٧٥	٧٤	٨٣	٩٢
٥٤	٤٥	٦٤	٦٠	٦٨	٦٦

والمطلوب :

وضع تلك الدرجات فى جدول تكرارى مزدوج.

٦- فيما يلى أرقام فرضية عن عدد الأسر المستفيدة من الوحدات الصحية ببعض القرى والمطلوب عرضها بيانياً.

البيان	١٩٧٠	١٩٨٠
أ	٢٥٧١٠	٥٢٦٧٥
ب	٤٠٧٦	٩٨٨١
جـ	٥٩٥٨	٩٥١٥
د	١٥٣٤	٧٧٢١
هـ	٢٩٨٢	٣٧٨٩

٧ - فيما يلى عدد السكان فى مدينة ما موزعة ذكوراً وأناثاً والمطلوب عرضها بيانياً بالأساليب التى تراها مناسبة:

السنة	ذكــور	إناث	الجملة
١٩٨٠	١٧٩٣٨	٦٣٦٢	٢٤٣٠٠
١٩٨١	١٥٦١١	٥٧٠٠	٢١٣١١
١٩٨٢	١٤٨٤٢	٥٨٨٧	٢٠٧٢٩
١٩٨٣	١٥٢٤٥	٦٨٠١	٢٢٠٤٦

٨- فيما يلى درجات لعينة من الطلاب فى إحدى الامتحانات:

٩٩	٨٨	٨٢	٧٧	٧٤	٦٩
٩٦	٨٧	٨٠	٧٧	٧٣	٦٩
٩٣	٨٦	٨٠	٧٦	٧٣	٦٧
٩١	٨٤	٧٨	٧٤	٧١	٦٦
٩٠	٨٣	٧٨	٧٤	٧٠	٦٤
٩٠	٨٢	٧٧	٧٤	٧٠	٦١

(أ) ابدأ بعمل الفئة ٦٠ – ٦٢ واستكمل الفئات لتشمل جميع الدرجات ثم أرسم هيستوجراماً لتلك البيانات.

(ب) استخدم نفس الفئات لرسم المضلع التكرارى.

(جـ) أوصف توزيع تلك القيم.

(د) صنف الدرجات السابقة فى جدول تكرارى مدى كل فئة فيه خمس درجات.

٩ – تمثل البيانات التالية الحالة الزواجية لعينة ، تم سحبها من أحد الاحياء السكنية بمدينة القاهرة. والمطلوب تنظيم هذه البيانات، وعمل جدول توزيع تكرارى للحالة الزواجية لافراد العينة.

متزوج	مطلق	متزوج	متزوج	أرمل	متزوج
ارمل	متزوج	متزوج	كتزوج	متزوج	متزوج
متزوج	متزوج	متزوج	مطلق	أعزب	متزوج
متزوج	أعزب	متزوج	أعزب	أعزب	متزوج
متزوج	أعزب	أعزب	أعزب	أعزب	أعزب
ارمل	أعزب	أعزب	أعزب	أعزب	مطلق
أعزب	مطلق	أعزب	متزوج	أعزب	اعزب
مطلق	أرمل	أرمل	متزوج	أعزب	اعزب
ارمل	متزوج	متزوج	متزوج	متزوج	متزوج
متزوج	متزوج	متزوج	متـزوج	متزوج	متزوج

١٠-أكمل مايأتى :

(أ) يفضل استخدام النسب المئوية أو النسب بدلا من التكرارت فى عمل................

(ب) تعتبر الأعمده البيانية والدوائر بقطاعاتها أفضل اساليب الرسومات البيانية للمتغيرات المقاسة عند مستوى..................

(جـ) يعتبر المضلع شكلا بيانيا ملائما لمستوى............من المتغيرات، ويكون وضع القيم على المحور السينى (الافقى) باستخدام..............للتوزيع التكرارى.

(د) يستخدم المضلع التكرارى نقطة عند وسط كل فئة بدلا من ....لتمثيل التكرارات.

١١- فيما يلى عدد من الاختيارات احداها أو اكثر يكون الاجابة الصحيحة التى تكمل العبارة فى السؤال ذاتة. والمطلوب وضع علامة دائرة (o) على رقم الاختيار الصحيح.

أ - عندما يتم ضرب اى نسبة a proportion فى رقم (١٠٠) يكون الناتج:

(١) معدل

(٢) نسبة مئوية Percentage

(٣) نسبة Ratio

(٤) نسبة تراكمية Cumulative proportion

(٥) لاشىء له معنى

ب - غالبا ما تستخدم الأعمده البيانية فى التمثيل البيانى للبيانات المقاسة عند:

(١) مستويات اسمية ومستويات نسبة - فئة

(٢) مستويات اسمية ومستويات اعتيادية

(٣) مستويات اعتيادية ومستويات نسبة – فئة

(٤) لا لجميع المستويات المذكورة سابقا

(٥) جميع المستويات المذكورة سابقا فى ١، ٢، ٣

جـ- اذا قمت برسم خطوط مستقيمة تربط بين منتصف الأعمدة للمضلع، سوف تحصل على:

(أ) منحنى اعتيادى.

(ب) توزيع تكرارى.

(جـ) مضلع تكرارى.

(د) ستحدث مشكلة كبرى.

(د) عند عمل توزيع تكرارى فإن أول خطوة يجب عملها هى:

(١) ايجاد مدى القيم

(٢) أخذ الحدود الحقيقية

(٣) تحديد مدى الفئة

(٤) حساب الفئة المئوية للتوزيع

(٥) لا أجابة من الاجابات الاربع السابقة.

(هـ) أن النقاط التى أقوم بالتوصيل بينها بخطوط فى المضلع التكرارى تتطابق مع:

(أ) الحد الأدنى للفئة.

(ب) الحد الاعلى للفئة.

(جـ) النقطة المتوسطة للفئة.

(د) النقاط النهائية للفئة.

(هـ) ليس للفئة أى حدود.

الفصل الرابع

مقاييس النزعة المركزية

مقدمـــــــة

- المـــنوا ل
- الوسيط
- المتوسط الحسابى
- العلاقة بين المتوسطات الثلاثة
- المتوسط المرجح



الفصل الرابع

مقاييس النزعة المركزية

مقدمة :

تناولنا فى الفصل السابق استخدام الأشكال والرسومات البيانية للتوزيعات التكرارية كإحدى الطرق الإحصائية فى تنظيم وتلخيص وتبسيط البيانات من خلال علاقة بيانية ذات معنى واضح. أيضاً تناولنا أنواع الجداول التكرارية فى تلخيص تلك البيانات التى تمثل قيماً لمتغير معلوم.

بالإضافة إلى الطريقتين السابقتين، يوجد العديد من الطرق أو الأساليب الإحصائية التى تزيد من تفصيلات وصف التوزيعات. ويستلزم لوصف التوزيع، توافر ثلاثة أنواع من المعلومات هى:

(أ) معرفة شكل التوزيع.

(ب) مؤشرات دالة على مواضع تلك التوزيعات على مقياس إحصائى مناسب.

(جـ) مؤشرات أو دلالات عن مدى التباينات بين القيم التى يتخذها المتغير فى كل موضع Location داخل التوزيع.

لمعرفة أى توزيع يجب على الدارس أن يبحث عن إجابات للبنود الثلاثة السابقة التى يمكن صياغتها فى ثلاثة استفسارات على النحو التالى: ما شكل التوزيع ؟ أين تكون مواضع قيم التوزيع ؟ إلى أى مدى تتصف قيم التوزيع بخاصية الانتشار؟

فيما يتعلق بالإجابة عن السؤال الأول، سبق مناقشة الأشكال المختلفة للتوزيعات وطرق تمثيلها بيانياً وبالرسومات التى من بينها المدرج التكرارى والمضلع التكرارى. وبالنسبة للسؤال الثانى، فإن أكثر الطرق الإحصائية شيوعاً فى قياس الموضع، هى مقاييس النزعة المركزية Measures of central Tendency.

وتشمل مقاييس النزعة المركزية المتوسطات التى تنزع للتمركز نحو منتصف التوزيع . وهى المنوال Mode والمتوسط الحسابى Arithmetic Mean والوسيط Median.

أما إجابة السؤال الثالث فتختص بخاصية تشتت Dispersion or Variability القيم فى التوزيعات. ومن أكثر مقاييس التشتت شيوعاً، التباين Variance، الانحراف المعيارى Standard Deviation ، الانحراف الربيعى ، الانحراف المتوسط Mean Deviation و المدى Range.

ترجع أهمية ارتباط النزعة المركزية بالتشتت إلى أن المتوسطات التى تستخدم فى قياس النزعة المركزية تصف المجموعة بقيمة واحدة بدلاً من استخدام قيم كثيرة تكون هذه المجموعة. حيث أن استخدام قيمة متوسطة كالمتوسط الحسابى لا تستطيع أن تعطى وصفاً كاملاً لتوزيع المجموعة أو أن تستخدم فى إجراء مقارنات بين قيم المجموعة ومجموعات أخرى.

من ثم كان من الضرورى لاستكمال الصورة الوصفية لتوزيع قيم متغير ما، أن نبرز مدى تباعد القيم بعضها عن بعض أو تجمعها Clustering باستخدام مقاييس التشتت المختلفة . وتبرز أهمية معرفة تشتت القيم لمتغير ما أو خاصية معينة فى وضع الخطط المناسبة أو الطرق الصحيحة للتعامل معها. فمثلاً لو تبين لمدرب إحدى فرق كرة القدم أن فريق الخصم يعتمد فى معظم حالات الهجوم على الجانب الأيمن فى تسجيل الأهداف أو نقل الكرة فإن ذلك يجعله يراعى عند وضع خطة اللعب أن يقوى الجهة اليسرى الدفاعية من فريقه وبالمثل لو تبين لأستاذ مادة الإحصاء من خلال الاختبارات الدورية لطلابه خلال العام الدراسى أن هناك تفاوتاً فى الدرجات التى حصلوا عليها. فإن ذلك يجعله يصيغ أسئلة نهاية العام بحيث يمكن لجميع الطلبة أن تجيب على قدر من الأسئلة أو كلها.

مقاييس النزعة المركزية :

لما كانت خاصية الوضع Location فى التوزيعات تتطلب الترتيب فإن تلك الخاصية تصبح غير ذات قيمة أو معنى للبيانات الكيفية أو التصنيفية Nominal Data ومن ثم فإن مقاييس النزعة تلائم فقط القياسات على المقاييس الترتيبية

Ordinal، الفئوية Interval وأيضاً مقاييس النسبة Ratio Scales فمثلاً يمكن استخدام الوسيط فى حالة البيانات الترتيبية، كما يتطلب استخدام المتوسط مقياس فئوى. كما يعتبر المنوال أفضل المقاييس لتمثيل البيانات النوعية.

ويجدر التنويه هنا إلى أهمية المقاييس الثلاثة للنزعة المركزية (المنوال. المتوسط الحسابى، الوسيط) فى مجال الاستنتاج الإحصائى Statistical Inference الذى يهتم أساساً بمعرفة طبيعة توزيع ظاهرة مافى المجتمع الأصلى والتوصل من خلال تصميم العينات إلى تعميمات توصف هذا المجتمع. ومن هنا يتضح أهمية المقاييس الثلاثة والتى يتم حسابها من العينة بوصفهم تقديرات للقيم الحقيقية المتوسطة للمجتمع الأصلى (M). ومن ثم يمكن أن تسمى تلك المقاييس بالتقديرات (Estimates) للمؤشرات أو المعاملات Parameters للمجتمع الأصلى.

المـــنوال :

يعتبر المنوال Mode أهم مقاييس النزعة المركزية بل هو المقياس الوحيد الذى يستخدم لحساب المتوسط لظاهرة لا يمكن قياسها بالمقياس الكمى مثل متغيرات المهنة ، النوع Sex، اللون فكل منها تمثل ظاهرة أو خاصية وصفية. والمنوال للبيانات الرقمية المتقطعة (Discrete) أو للبيانات الوصفية عبارة عن قيمة هذا المتغير الذى يحصل على أعلى التكرارات.

مثال (١) :

الجدول التالى يوضح النسبة المئوية لتوزيع تكرارات لخمس فئات حرفية فى المجتمع الريفى . والمطلوب تحديد الفئة المنوالية بين تلك الفئات على ضوء التعريف السابق للمنوال .

م	الفئات المهنية	٪
١	مزارعون	٤٥ ٪
٢	تجار ماشية	٢٥ ٪
٣	حلاق صحة	١٥ ٪
٤	بناء	١٠ ٪
٥	قابلة	٥ ٪
		١٠٠ ٪

من التكرارت الموضحة بالجدول ، يتضح أن فئة المزارعين تمثل الفئة المنوالية لأن القيمة التكرارية عندها (٤٥ ٪) هى الأعلى قياساً بباقى التكرارت.

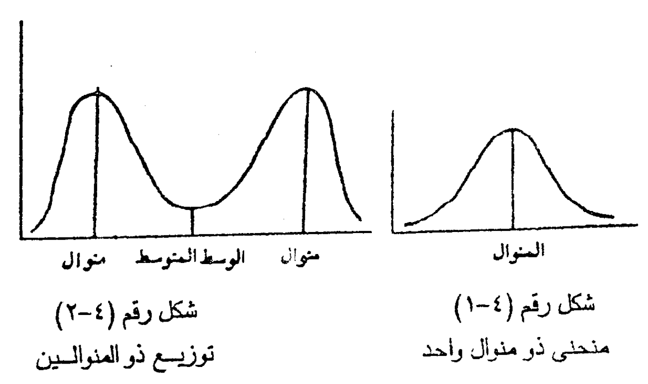
المنوال والبيانات المتصلة Mode and the Continuous Data :

من خلال التعريف السابق للمتغير المتصل والذى تناولناه فى الفصل الثانى لا يوجد منوال أو فئة منوالية لمتغير من هذا النوع مثل الطول، الوزن وغيرهما حيث تختلف المقاييس والأوزان وحيث نجد تكراراً واحداً فقط لكل قيمة تجريبية لهذا المتغير سواء كان القياس بالمتر أو البوصة فى حالة الطول أو بالكيلو جرام والرطل فى حالة الوزن. بالمثل لا يوجد منوال أو فئة منوالية إذا كانت البيانات عبارة عن مفردات كما هو الحال فى مسابقات العاب القوى فمثلاً لو حصل متسابق على درجات فى بعض اللعبات وكانت ٤١,٣٥ ، ٧٨,٥٦ فمثل هذه السلسلة من الأرقام لا يوجد لها منوال أو فئة منوالية.

من جهة أخرى، إذا أمكن فى حالة قياس متغير متصل مثل الوزن أو الدخل مثلاً بحيث يمكن تجميع البيانات التجريبية فى شكل مجموعات أو فئات. فيمكن بواسطة دلالة التكرارات لتلك المجموعة أو الفئات أن نحصل على المنوال أو الفئة المنوالية والتى عندها تحدث أعلى قيمة تكرارية وتعبر القيمة المتوسطة التى يتحقق عندها هذا الشرط (أعلى تكرار).

هذا ويجدر الإشارة إلى احتمالية وجود أكثر من فئة منوالية فى هذا النوع من البيانات المتصلة. أو بمعنى آخر، قد لايكون للمنوال قيمة واحدة فقد توجد قيمة

منوالية أخرى ويطلق على التوزيع فى هذه الحالة التوزيع ذو المنوالين Bimodal كما يطلق على التوزيع وحيد القيمة للمنوال، التوزيع ذو المنوال الواحد Unimodal ويكون شكل التوزيع فى كل حالة منهما كما يتضح فى الشكلين رقمَ (٤-١) ، (٤-٢).



شكل رقم (٤-٢)
توزيـع ذو المنوالــين

شكل رقم (٤-١)
منحنى ذو منوال واحد

هذا وقد نتوقع وجود أكثر من فئة منوالية كلما كانت التكرارات مذبذبة بين الزيادة والنقصان ثم الزيادة وهكذا.

مثال (٢) :

أحسب المنوال من البيانات التكرارية التالية وذلك للدخل السنوى لعدد (٥٠) أسرة حضرية:

الدخل السنوى بالجنيه	ك	النقطة المتوسطة
صفر - ١٩٩٩	٥	١٠٠٠
٢٠٠٠ - ٣٩٩٩	٢٢	٣٠٠٠
٤٠٠٠ - ٥٩٩٩	١٥	٥٠٠٠
٦٠٠٠ - ٧٩٩٩	٨	٧٠٠٠
مجـ	٥٠	

فى هذا المثال، نجد أن المنوال عند (٣٠٠٠) قيمة متوسطة تناظر أكبر التكرارات (٢٢).

مثال (٣) :

أوجد قيمة المنوال للارقام العددية التالية:

٣٩ ، ٢٥ ، ١٤ ، ٣٩ ، ٢٥ ، ٢٠

الحـــل :

١- رتب القيم ترتيباً تصاعدياً (الأصغر فالأكبر وهكذا...)

٢- حدد القيمة أوالقيم الأكثر شيوعاً.

١٤ ، ٢٠ ، ٢٥ ، ٢٥ ، ٣٩ ، ٣٩.

فى هذا المثال، نجد قيمتين منواليتين وفقاً لتكرارهما القيمة المنوالية الأولى هى (٢٥) والقيمة المنوالية الثانية هى (٣٩).

نخلص مما سبق إلى تعريف عام للمنوال بأنه القيمة الأكثر شيوعاً أو انتشاراً. ومن ثم يتوقف تحديد قيمة المنوال على تكرار القيم داخل المجموعة.

طرق حساب المنوال من البيانات المبوبة:

يمكن تقدير قيمة المنوال وتحديد الفئة المنوالية، أما بالطرق الحسابية باستخدام المعادلات وأما بطريقة الرسومات البيانية وأيضاً بكليهما معاً وتوجد خمسة طرق لتقدير المنوال منها ثلاث حسابية وطريقتان بالرسم والطرق الحسابية الثلاث هم :

(أ) طريقة مركز الفئة المنوالية.

(ب) طريقة بيرسون (الفروق الدقيقة).

(جـ) طريقة العزوم الرياضية.

وسوف نكتفى فقط بالطريقتين (أ) ،(ب) فى هذا الفصل.

تقدير المنوال باستخدام طريقة مركز الفئة المنوالية :

تعتبر هذه الطريقة أبسط الطرق الثلاث الحسابية وأقلهم دقة ، نظراً لأن المنوال عادة ما ينحاز إما صوب بداية الفئة المنوالية أو ناحية نهايتها تبعاً لتكرارات الفئتين السابقة واللاحقة للفئة المنوالية.

ومن ثم لا نتوقع تطابق قيمة المنوال مع مركز الفئة المنوالية إلا إذا تساوى تكرارى الفئتين السابقة واللاحقة للفئة المنوالية.

مثال (٤) :

احسب المنوال من التوزيع التكرارى للعمر لعينة من العاملين ، وذلك باستخدام مركز الفئة المنوالية.

فئة السن	ك
١٥ –	١٥
٢٥ –	٣٥
٣٥ –	٢٥
٤٥ –	١٥
٥٥ –	٦
٦٥ –	٤

فنجد أن الفئة المنوالية (٢٥–) هى الفئة المنوالية

$$\text{مركز الفئة المنوالية} = \frac{٢٥ + ٣٥}{٢} = ٣٠$$

قيمة المنوال = ٣٠

فى هذه الطريقة، إذا أمكن للباحث معرفة القيمة المنوالية لأى توزيع تكرارى أو بمعنى آخر معرفة الفئة التى تناظر أكبر التكرارات فإن مركز الفئة يكون المنوال. ويمكن حساب قيمة مركز الفئة المنوالية بجمع قيمتى البداية والنهاية لتلك الفئة ثم قسمتها على (٢) . كما يتضح ذلك من المثال السابق، والذى تضمن

توزيعات تكرارية مكونة من (١٠٠) فرد وفقاً لفئة السن تم سحبها من كشوف العضوية لأحد الأندية الرياضية :

طريقة بيرسون (الفروق الدقيقة) :

تماثل طريقة كارل بيرسون الطريقة التالية فى حساب قيمة المنوال باستخدام تكرارى الفئتين السابقة (ك١) واللاحقة (ك٢) للفئة المنوالية مع اختلاف واحد هو استخدام بيرسون للفروق بين تكرارى الفئتين بالنسبة لتكرار الفئة المنوالية.

ففى طريقة حساب المنوال باستخدام تكرارى الفئتين السابقة واللاحقة للفئة المنوالية، يتم تقسيم مدى الفئة المنوالية تقسيماً يتناسب مع التكرارين السابق واللاحق فإذا كان تكرار الفئة اللاحقة أكبر من تكرار الفئة المنوالية ينحرف المنوال ناحية القيم الكبرى والعكس صحيح. ،من ثم يمكن حساب قيمة المنوال باستخدام المعادلة الآتية :

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{تكرار الفئة بعد المنوالية}}{\text{تكرار الفئة السابقة للمنوالية} + \text{تكرار الفئة اللاحقة}} \times \text{مدى الفئة المنوالية}$$

مثال (٥) :

أحسب المنوال من التوزيع التكرارى لدخول عينة تتكون من ٢٥ عاملاً وذلك من الجدول التالى باستخدام كل من طريقة تكرارى الفئتين (السابقة واللاحقة) ، وطريقة بيرسون:

فئات الدخل	ك
١٥ –	٥
٢٠ –	١٢
٢٥ –	٤
٣٠ – ٣٥	٤
مجـ	٢٥

وحيث طول الفئة = ٥

الحد الأدنى للفئة المنوالية = ٢٠

تكرار الفئة قبل المنوالية = ٥

تكرار الفئة بعد المنوالية = ٤

$$\text{المنوال} = ٢٠ + \frac{٤}{٤ + ٥} \times ٥$$

= ٢٢,٢

الحـل : باستخدام طريقة بيرسون.

خطـوات الحــل :

(أ) يتم عمل جدول يضم الفئة المنوالية والفئتين السابقة واللاحقة لها بالفئات والتكرارت مع إضافة عمود للفروق.

(ب) يحسب الفرق الأول بين تكرارى الفئة المنوالية والفئة السابقة لها ثم يحسب الفرق الثانى بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

(جـ) يستخدم نفس العلاقة السابقة فى حساب المنوال بطريقة تكرار الفئتين مع استبدال التكرارات فى المقدار الكسرى بالفروق وفق نفس الترتيب فتصبح العلاقة لحساب المنوال والمعروفة بمعادلة بيرسون على النحو التالى:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{الفرق بين تكرارى المنوال والسابق}}{\text{الفرق بين تكرارى المنوال والسابق} + \text{الفرق بين تكرارى المنوال واللاحق}} \times \text{طول الفئة}$$

ويرمز للفرق بين تكرارى المنوال والسابقة بالرمز(Δ١) ويرمز للفرق بين تكرارى المنوال واللاحق بالرمز(Δ٢) الحد الادنى للفئة المنوالية بالحرف (ح د) وطول الفئة بالحرف (ط) .

فئات	ك	فروق
١٥ – ٢٠	٥	
٢٠ – ٢٥	١٢	٧ Δ_1
٢٥ – ٣٠	٤	٨ Δ_2

$$\therefore \text{المنوال} = ح_د + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times ط$$

$$\text{المنوال} = ٢٠ + \frac{٧}{٧ + ٨} \times ٥$$

$$= ٢٠ + ٢,٣ = \underline{\overline{٢٢,٣}}$$

وهى تقريباً نفس القيمة السابقة مع قدر يسير من الدقة.

ولعل الاختلاف البسيط بين قيمتى المنوال يدل على أن المنوال مقياس غير مستقر – وكما سيتضح فيما بعد – نجد أن قيمته تتوقف على تبويب البيانات فى حالة التوزيعات التكرارية، فلو كان التبويب متصفاً باختلاف أطوال الفئات لاختلفت تبعاً لذلك قيمة المنوال.

ففى المثال السابق راعينا أن تكون أطوال الفئات متساوية ، ولكن هناك حالات لاتتحقق فيها هذه الخاصية. وفى مثل تلك الحالات، لا يستخدم الباحث علاقة بيرسون أو أى من العلاقات السابقة.

حساب المنوال إذا كانت أطوال الفئات غير متساوية :

فى هذه الحالة يقوم الباحث بإجراء تعديلات فى التكرارات المدونة بجدول البيانات وذلك بأن يقسم كل تكرار على طول الفئة المناظرة لهذا التكرار. وفى هذه الحالة يمكن للباحث أن يستخدم معادلة بيرسون السابقة.

أما إذا أراد الباحث أن يستخدم التكرارت الأصلية دون إدخال أى تعديلات فيمكن أن يحسب المنوال فى هذه الحالة من المعادلة التالية:

المنوال = الحد الأدنى للفئة المنوالية +

$$\left[\frac{\text{طول الفئة قبل المنوالية + تكرار الفئة المنوالية}}{\text{طول الفئة بعد المنوالية × تكرار الفئة قبل المنوالية + طول الفئة قبل المنوالية × تكرار الفئة بعد المنوالية}} \right]$$

× مدى الفئة بعد المنوالية

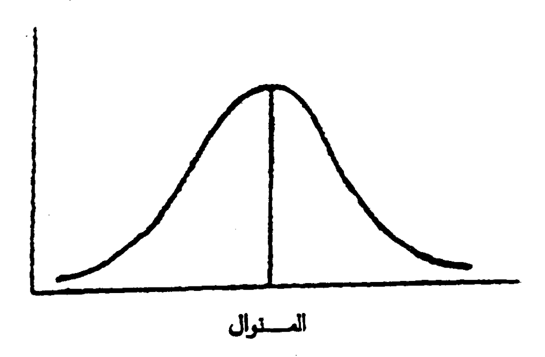
حساب المنوال من الرسومات البيانية:

(أ) من المنحنى التكرارى Frequency Curve

من التعريف العام للمنوال بأنه دائماً القيمة التى تقابل أكبر تكرار. نقول أننا إذا رسمنا التوزيع التكرارى وقمنا بإسقاط عمود من أعلى نقطة فى المنحنى التكرارى (التى تمثل أعلى قيمة تكرارية) فإنه سوف يقطع المحور الأفقى أو السينى فى نقطة هى المنوال. كما يتضح ذلك من الشكل رقم (٤-٣) .

(ب) من المدرج التكرارى:

يكتفى عند رسم المدرج التكرارى لحساب المنوال باختيار ثلاثة مستطيلات فقط يمثل الأوسط الفئة المنوالية وعلى كل جانب منه يقوم الباحث برسم:



شكل رقم (٤ – ٣)

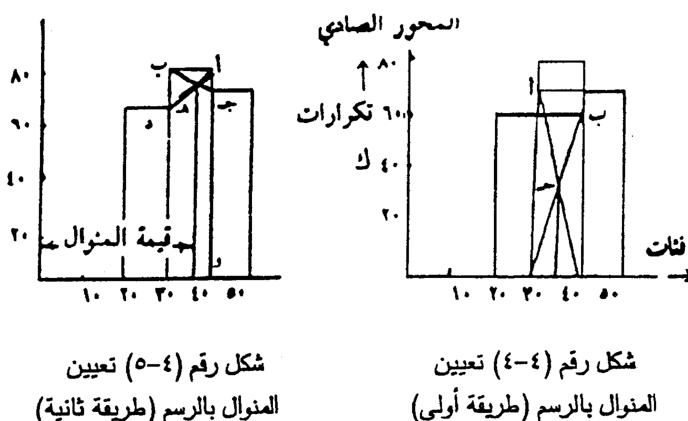
مستطيل يمثل القيم التكرارية للفئة السابقة مباشرة للفئة المنوالية فى ترتيب الجدول ويكرر نفس العمل على الجهة الأخرى من مستطيل الفئة المنوالية حيث يقوم برسم مستطيل بقيم تكرارية ومدى الفئة للفئة التالية مباشرة للفئة المنوالية. ولايجاد قيمة المنوال هناك طريقتان كما يتضح ذلك من الشكلين (٤-٤) ، (٤-٥) للمدرج التكرارى:

الطريقة الأولى :

يقوم الباحث بعد رسم المستطيلات الثلاثة، بمد خطى قمة كل مستطيل على جانبى المستطيل المنوالى حتى يقطعا الخطين الرأسيين لحدود المستطيل المنوالى عند نقطتين أ، ب مثلاً، بعد ذلك يقوم الباحث برسم خط قطرى أو محورى من النقطة (أ) حتى تلاقى أدنى نقطة التقاء مقابلة فى الضلع الرأسى الأول للمستطيل المنوالى. ويكرر نفس العمل بالنسبة للنقطة (ب) فيرسم خط قطرى من هذه النقطة حتى أدنى نقطة التقاء الضلع الرأسى الثانى للمستطيل المنوالى مع المحور الأفقى ونتيجة لذلك يلتقى القطران أو المحوران فى نقطة تقاطع (جـ) داخل مساحة المستطيل المنوالى. يقوم الباحث بعد ذلك بإسقاط عمود من النقطة (جـ) ليقابل المحور الأفقى فى نقطة (د) وعندئذ يكون البعد السينى لهذه النقطة (د) هو قيمة المنوال.

الطريقة الثانية:

لو افترضنا أن خط القمة لتكرارت المستطيل المنوالى هو أ، ب وأن مستطيل الفئة السابقة يشترك مع المستطيل المنوالى فى ضلع واحد ولكن بطول يبدأ من المحور السينى حتى النقطة (جـ). أيضاً أن مستطيل الفئة اللاحقة يشترك مع المستطيل المنوالى فى ضلع واحد من الجهة المقابلة ولكن بطول يبدأ من المحور السينى وبقيمة تكرارات حتىَ النقطة (د) كما يتضح ذلك من الشكل رقم (٤-٥) فلو قام الباحث برسم خط محورى يصل النقطتين أ، د ثم يرسم خط محورى آخر يصل النقطتين ب، جـ فإن المحورين يتقاطعان عند نقطة ولتكن (هـ) ليقطع المحور السينى عند نقطة (و) فإن البعد (الإحداثى) السينى لهذه النقطة يمثل قيمة المنوال.



شكل رقم (٤-٤) تعيين المنوال بالرسم (طريقة أولى)

شكل رقم (٤-٥) تعيين المنوال بالرسم (طريقة ثانية)

المتوسط الحسابي :

حساب المتوسط الحسابي من البيانات غير المبوبة :

يعتبر المتوسط الحسابي والوسيط أهم مقاييس النزعة المركزية استخداماً في البحوث الاجتماعية وإن كان المتوسط الحسابي أكثرهما شيوعاً. ولعل السبب في ذلك يرجع إلى أن المتوسط الحسابي هو أصدق المقاييس الثلاثة تمثيلاً للمجموعة تحت الدراسة ، وبالتالي تصبح تلك المجموعة مؤشراً جيداً للمجموعة الأصلية نظراً لأن المتوسط دائماً يكون من نفس وحدات المتغير. ومن تلك الخاصية يعرف المتوسط بأنه حاصل قسمة مجموع القيم على المجموع الكلي لعدد الحالات. ومن ثم لو افترضنا أن القيم هي $س_1$ ،$س_2$ ،$س_3$... حتى $س_ن$ وإن عدد الحالات هي (ن) فإن قيمة المتوسط الحسابي ويرمز له ($\bar{س}$) يتم حسابها من المعادلة التالية:

$$\bar{س} = \frac{س_1 + س_2 + س_3 + \cdots\cdots + س_ن}{ن} \qquad \therefore \bar{س} = \frac{مجـ\ س}{ن}$$

حيث $س_1$ تمثل قيمة المفردة الأولى، $س_2$ تمثل قيمة المفردة الثانية وهكذا حتى $س_ن$ قيمة المفردة الأخيرة.

مثال (٦) :

ذا كانت ١٢، ١٨ ، ١٠ ، ٧ هى الدرجات التى حصلت عليها أربع طالبات فى اختبار نصف العام لمادة الإحصاء الاجتماعى. أحسب المتوسط الحسابى لتلك القيم؟

$$\therefore \bar{س} = \frac{١٢ + ١٨ + ١٠ + ٧}{٤} = ١١,٧٥$$

ويتصف المتوسط الحسابى بخاصية جبرية أساسية وهى أن مجموع انحرافات القيم عن المتوسط الحسابى لها لا بد أن يساوى صفرا. أى أن:

$$مجـ (س - \bar{س}) = صفراً$$

ولا غرابة فى اتصاف المتوسط الحسابى بتلك الخاصية والتى يمكن استنباطها منطقياً من التعريف السابق كما يمكن إثباتها من المثال التالى:

مثال (٧) :

نفرض أننا نريد حساب المتوسط من الأرقام التالية:

٧٢ ، ٨١ ، ٨٦ ، ٦٩ ، ٥٧.

الحـــل :

$$\text{من العلاقة السابقة } \bar{س} = \frac{٧٢ + ٨١ + ٨٦ + ٦٩ + ٥٧}{٥} = ٧٣$$

فلو قمنا بطرح قيمة سَ وهى (٧٣) من كل قيمة من القيم الست سيكون ناتج جمع الفروق = صفراً كما يتضح ذلك من الجدول التالى:

س	س – ٧٣	س – ٧٠
٧٢	– ١	٢
٨١	٨	١١
٨٦	١٣	١٦
٦٩	– ٤	–١
٥٧	–١٦	– ١٣
مجـ	صفر	١٥

من جهة أخرى، لو افترضنا قيمة معينة للمتوسط الحسابى ولتكن مثلاً (٧٠) فى المثال السابق فهل يمكن التوصل للقيمة الحقيقية للمتوسط الحسابى فى هذا المثال؟ نعم يمكن التوصل للقيمة الحقيقة للمتوسط الحسابى بأن نبدأ أولاً بالتأكد من أن جميع الفروق بين س وقيمة س َ، تساوى صفراً فى هذه الحالة يكون المتوسط الفرضى هو المتوسط الحقيقى . أما إذا كان مجموع الفروق لا يساوى صفراً فأما أن تكون قيمة مجموع الفروق موجبة أو سالبة. ففى المثال السابق نجد أن قيمة مجموع الفروق =+١٥ . وهذا يعنى أن المتوسط المفترض يقل عن القيمة الحقيقية للمتوسط وهو (٧٣) من ثم نقوم بقسمة مجموع الفروق على عدد الحالات وهى (٥) فى هذا المثال فيكون ناتج القسمة هو (٣). نقوم بعد ذلك بإضافة هذا الناتج إلى المتوسط الفرضى فنحصل على نفس القيمة للمتوسط الحقيقى ٧٠ + ٣ = ٧٣ . أما إذا كانت قيمة مجموع الفروق سالبة فإننا نطرح ناتج القسمة (بدلاً من إضافته) من المتوسط الفرضى.

نخلص من ذلك إلى العلاقة التالية التى تربط بين المتوسطين الحقيقى والفرضى:

$$\text{المتوسط الحقيقى} = \text{المتوسط الفرضى} \pm \frac{\text{مجموع الانحرافات (الفروق) عن المتوسط الفرضى}}{\text{عدد الحالات}}$$

من هذه العلاقة يمكن استنتاج خاصية هامة وهى أنه كلما كان المتوسط الفرضى أقرب من قيمة المتوسط الحقيقى، قلت قيمة مجموع الفروق.والخاصية الثانية للمتوسط الحسابى والتى تستخدم فى قياس التباين الكلى هى أن مجموعة مربع الانحرافات لكل قيمة عن المتوسط الحسابى دائماً أقل من مجموع مربع الإنحرافات لهذه القيمة عن أى قيمة أخرى غير المتوسط الحسابى. أما الخاصية الثالثة للمتوسط الحسابى فهى أن قيمته دائماً تقع بين أعلى القيم وأدناها ولا يأتى وقوعه دائماً فى منتصف القيم، مع التقيد فى الوقت ذاته بالخاصيتين السابقتين. وهذه الخصائص الثلاث يجب أن تتمثل فى قيمة المتوسط الحسابى مهما زاد أو قل عدد القيم أو حجم البيانات.

حساب المتوسط الحسابى من البيانات المبوبة Grouped Data :

مثال (٨) :

فيما يلى توزيع تكرارى لدرجات عينة من الطلاب فى امتحان مادة الإحصاء والمطلوب حساب المتوسط الحسابى لهذه المجموعة.

فئات	ك
٢٠ –	١٢
٣٠ –	٨
٤٠ –	٢
٥٠	٩
٦٠ –	٥٠
٧٠ –	١٣
٨٠ –	٧
٩٠ – ١٠٠	٨
مجـ	١٠٩

خطوات الحل :

١ – أوجد مركز كل فئة (س).

٢ – ضرب مركز الفئة فى التكرار المناظر للفئة ذاتها (س ك) .

٣ – تطبيق المعادلة الآتية:

$$\text{س}' = \frac{\text{مجـ س} \times \text{ك}}{\text{مجـ ك}}$$

ف	ك	س	س ك
٢٠ -	١٢	٢٥	٣٠٠
٣٠ -	٨	٣٥	٢٨٠
٤٠ -	٢	٤٥	٩٠
٥٠ -	٩	٥٥	٤٩٥
٦٠ -	٥٠	٦٥	٣٢٥٠
٧٠ -	١٣	٧٥	٩٧٥
٨٠ -	٧	٨٥	٥٩٥
٩٠ - ١٠٠	٨	٩٥	٧٦٠
	١٠٩		٦٧٤٥

نخلص مما سبق إلى أنه فى حالة (المتغير الواحد) فى البيانات التكرارية المجدولة مثل الطول، الأجرأو الوزن فإن قيمة المتوسط الحسابى (س) تحسب من العلاقة التالية:

$$\text{سَ} = \frac{\text{مجـ س × ك}}{\text{مجـ ك}}$$

$$= \frac{٦٧٤٥}{١٠٩} = ٦١,٩$$

أيضاً فى حالة البيانات غير المبوبة مثل سلسلة من الأرقام كما ذكرنا فى مثال سابق أو تسجيل عدد من الأرقام فى إحدى التمارين الرياضية يحسب المتوسط الحسابى (سَ) من العلاقة التالية:

$$\text{سَ} = \frac{\text{مجـ س}}{\text{ن}}$$

حساب المتوسط الحسابى من التوزيعات التكرارية باستخدام الوسط الفرضى:

إن المشكلة الكبرى التى تواجه الإحصائى فى جداول القيم التكرارية الموزعة إلى فئات هى عدم معرفة قيم الإفراد جميعها ففى فئات السن أو الطول

يصعب تحديد الطول الحقيقى لكل مفردة داخل الفئة. وإن كل ما يتوفر من بيانات هما حدى الفئة (البداية والنهاية) . ولسهولة حساب المتوسط الحسابى فى هذه الحالة يفترض الإحصائى قيماً متساوية لكل المفردات داخل الفئة قيمة متوسطة هى مركز الفئة. ويمكن الحصول على قيمة مركز الفئة أما بجمع حدى الفئة الواحدة وقسمة الناتج على (٢) أو بإضافة نصف مدى الفئة ذاتها إلى الحد الأدنى لها. ففى فئة طولية من ١٥٠ – ١٧٠ سم تكون قيمة مركز الفئة ١٦٠ سم.

وتتلخص عمليات حساب المتوسط الحسابى فى الخطوات التالية:

١ – يتم تحديد مركز الفئة.

٢ – يتم اختيار وسط فرضى (سً) يفضل أن يكون مركز الفئة لأكبر تكرار أو الأقرب لها فى القيمة.

٣ – يتم حساب انحرافات المركز الفئوية عن الوسط الفرضى (سً) فلو رمزنا للانحراف (ح) ∴ ح = (س – سً)

٤ –يتم ضرب الانحراف فى التكرار المناظر وذلك لكل فئة فنحصل على مجموع حاصل ضرب ح × ك.

٥ – يتم قسمة الناتج من مجموعة ح ك على مجموع التكرارات فنحصل على قيمة = $\frac{\text{مجـ ح ك}}{\text{مجـ ك}}$

٦ – تستخدم المعادلة السابقة لحساب المتوسط الحسابى بدلالة المتوسط الفرضى مع تغيير طفيف فى مقدار الفروق (المقدار الثانى من الطرف الأيسر لنفس المعادلة) بحيث تصبح المعادلة كالآتى:

$$\text{سَ} = \text{سً} \pm \frac{\text{مجـ ح كَ}}{\text{مجـ ك}}$$

مـــثال (٩) :

يوضح الجدول التكرارى التالى توزيع الأجور الشهرية لعدد (٣٥) عاملاً داخل إحدى الشركات الصناعية. والمطلوب حساب المتوسط الحسابى باستخدام وسط فرضى .

فئات الأجر الشهرى بالجنية المصرى	التكرار (ك)	(مركز الفئة)	الانحراف (ح)	ح ك عند ٦٧
٤٩ -	٩	٥١	- ١٦	- ١٤٤
٥٣ -	٢٧	٥٥	- ١٢	- ٣٢٤
٥٧ -	٣٩	٥٩	-٨	- ٣١٢
٦١ -	٤٨	٦٣	-٤	- ١٩٢
٦٥ -	٦٠	٦٧	صفر	صفر
٦٩ -	٤٥	٧١	+ ٤	١٨٠
٧٣ -	٣٩	٧٥	+ ٨	٣١٢
٧٧ -	٢٤	٧٩	+ ١٢	٢٨٨
٨١ -	٩	٨٣	+ ١٦	١٤٤
المجموع	٣٠٠			- ٩٧٢
				+ ٩٢٤
				- ٤٨

$$\text{س}' = ٦٧ - \frac{٤٨}{٣٠٠} = ٦٧ - ٠,١٦$$

$$= ٦٦,٨٤$$

هذا ويمكن إضافة عمود آخر للجدول بإيجاد قيمة للانحرف الجديد (حَ) لتقليل العمليات الحسابية وذلك بقسمة قيمة ح لكل فئة على طول الفئة.

وفى هذه الحالة نستخدم معادلة حساب المتوسط الحسابى مع ضرب طول الفئة فى المقدار الكسرى الثانى من الطرف الأيسر لتلك المعادلة:

أى:

$$\text{س}^- = \text{س}' + \frac{\text{مجـ ح ك}}{\text{مجـ ك}} \times \text{طول الفئة}$$

الوســـيط :

يعتبر الوسيط مقياساً ترتيبياً Ordinal على عكس المتوسط الحسابى والذى لا يستخدم إلا كمقياس كمى Interval Measurement ويعرف الوسيط بأنه النقطة أو القيمة التى تقسم القيم التجريبية أو القياسات التى تسجل حول ظاهرة ما إلى مجموعتين ، شريطة أن يتساوى عدد القيم الأكبر عن الوسيط مع باقى القيم الأصغر منه والتى تليه فى الترتيب، حيث يتم ترتيب تلك القيم جميعها أما ترتيباً تصاعدياً أو تنازلياً . ولعل من تعريف الوسيط قد يبدو للقارىء لأول وهلة أن طريقة حسابه بسيطة وهذا الزعم ليس صحيحاً إلى حد كبير. ففى الواقع العملى يواجه الباحث عدد من الصعوبات فى كثير من الأحيان مثال ذلك حساب الوسيط للمتغيرات المتقطعة خاصة ، إذا كان عدد القيم زوجياً . كذلك قد يواجه الباحث بعض الصعاب فى حساب الوسيط لقيم متصلة خاصة إذا كان عدد تلك القيم صغيراً أو يتصف بكثير من التكرارات. ولعل ذلك من عيوب الوسيط حيث لا يفضل استخدامه كلما قل عدد القيم أو المشاهدات التجريبية .

ويتصف الوسيط بخاصية هامة وهى عدم تأثره بالقيم المتطرفة على جانبى منحنى التوزيعات لقيم مبوبة. ومن ثم يفضل فى تلك الحالات استخدام المتوسط الحسابى . ويمكن استخدام الوسيط فى حالة التوزيعات التكرارية التى تتصف بحدة الإلتواء وأيضاً فى حالة التوزيعات التكرارية المفتوحة . وهذه خاصية هامة يتميز بها الوسيط عن كل من المتوسط الحساب والمنوال معاً.

حساب الوسيط فى حالة البيانات غير المبوبة :

١ - إذا كان عدد القيم فردياً:

من التعريف السابق للوسيط، ففى حالة البيانات غير المبوبة، يتم أولاً ترتيبها تصاعدياً أو تنازلياً فإذا كان عدد تلك القيم فردياً تكون قيمة الوسيط هو حاصل قسمة هذا العدد إجمالاً مضافاً إليه الواحد الصحيح على (٢) . فلو فرضنا أن عدد القيم (ن) .

∴ الوسيط = $\frac{ن + ١}{٢}$ والناتج من المعادلة يمثل وضع القيمة الوسيطة داخل ترتيب كل القيم.

مثال (١٠) :

حصل ١٣ طالباً على الدرجات التالية فى اختبار أعمال السنة فى مادة الاجتماع الصناعى. أحسب الوسيط لتلك الدرجات:

٤ ، ٩ ، ١٥ ، ٨ ، ١٤ ، ٦ ، ٥ ، ٧ ، ٩ ، ١٢ ، ١٠ ، ١٣ ، ١٦ .

الحـــل :

يتم أولاً ترتيب الدرجات المعطاة وفقاً لقيمها ترتيباً تنازلياً أو تصاعدياً كما قلنا. وفى هذا المثال اخترنا الترتيب التصاعدى .

٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ٩ ، ١٠ ، ١٢ ، ١٣ ، ١٤ ، ١٥ ، ١٦ .

الوسيط = $\frac{١٣ + ١}{٢}$

= $\frac{١٤}{٢}$ = ٧ أى القيمة السابعة

أى أن الوسيط هو القيمة السابعة فى الترتيب التصاعدى وهى ٩ . حيث توجد (٦) قيم سابقة هى ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ وأيضاً ست قيم لاحقة للقيمة الوسيطية هى ١٠ ، ١٢ ، ١٣ ، ١٤ ، ١٥ ، ١٦ .

٢ - إذا كان عدد القيم زوجياً:

إذا كان عدد القيم زوجياً ، توجد قيمتين وسيطتين الأولى عند $\frac{(ن)}{٢}$ والثانية عند ($\frac{ن}{٢}$) + ١

مثال (١١) :

فيما يلى درجات عشرة طلاب والتى حصلوا عليها فى اختبار آخر العام لمادة النصوص الأجنبية فى قسم الاجتماع. والمطلوب حساب الوسيط.

٩٥ ، ٨٦ ، ٨٧ ، ٩٠ ، ٦٢ ، ٧٣ ، ٨٩ ، ٩٢ ، ٨٤ ، ٧٦.

الحـــل :

نرتب أولاً الدرجات السابقة ترتيباً تصاعدياً مثلاً فتكون:

٦٢ ، ٧٣ ، ٧٦ ، ٧٨ ، ٨٤ ، ٨٦ ، ٨٩ ، ٩٠ ، ٩٢ ، ٩٥

ترتيب الوسيط الأول = $\frac{ن}{٢}$

= $\frac{١٠}{٢}$ = ٥ أى القيمة الخامسة فى الترتيب

∴ قيمة الوسيط الأول = ٨٤

ترتيب الوسيط الثانى = ($\frac{ن}{٢}$) + ١

= $\frac{١٠}{٢}$ + ١ = ٦ القيمة السادسة فى الترتيب

∴ قيمة الوسيط الثانى هى ٨٦.

الوسيط = $\frac{\text{الوسيط الأول + الوسيط الثانى}}{٢}$

= $\frac{١٧٠}{٢}$ = ٨٥

حساب الوسيط من بيانات مبوبة (جداول تكرارية) :

فى حالة البيانات المبوبة، وبعد ترتيب القيم تصاعدياً أو تنازلياً، يقوم الدارس بعمل الجدول المتجمع الصاعد أو الهابط وفق الترتيب. ثم يحدد بعد ذلك ترتيب الوسيط من بين تلك القيم باستخدام العلاقة الآتية بفرض أن (ك) تمثل التكرار.

ترتيب الوسيط = $\frac{\text{مجـ ك}}{٢}$

والناتج من العلاقة السابقة يحدد مباشرة الفئة التى يقع بين حديها الأدنى والأعلى ترتيب الوسيط. ويطلق على تلك الفئة (الفئة الوسيطة).

الوســيط =

$$\text{الحد الأدنى للفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع السابق}}{\text{التكرار الأصلى للفئة}} \times \text{طول الفئة الوسيطية}$$

مثال (١٢) :

يوضح الجدول رقم (٤-١) توزيع درجات ١٠٠ طالب فى امتحان مادة الفلسفة المعاصرة. أوجد الوسيط من هذه البيانات المبوبة.

خطوات الحل

١ - تكوين الجدول المتجمع الصاعد أو الهابط.

٢ - حساب ترتيب الوسيط وهو يساوى $\frac{\text{مجـ ك}}{٢}$

٣ - تطبيق معادلة قيمة الوسيط.

جدول رقم (٤-١)

توزيع درجات الطلاب فى مادة الفلسفة المعاصرة

ف	ك
٤٠ -	٥
٥٠ -	٢٥
٦٠ -	٣٥
٧٠ -	٢٥
٨٠ - ٩٠	١٠
مجـ	١٠٠

حساب الوسيط بواسطة التكرار المتجمع الصاعد :

حدول رقم (٤-٢)

جدول التكرار المتجمع الصاعد

الحدود العليا	ك	تكرار متجمع الصاعد
أقل من ٥٠	٥	٥
أقل من ٦٠	٢٥	٣٠
أقل من ٧٠	٣٥	٦٥
أقل من ٨٠	٢٥	٩٠
أقل من ٩٠	١٠	١٠٠

ترتيب الوسيط = $\frac{\text{مجـ ك}}{٢}$

= $\frac{١٠٠}{٢}$ = ٥٠

الوسـيط = ٦٠ + $\frac{٥٠ - ٣٠}{٣٥}$ × ١٠

= ٦٠ + $\frac{٢٠٠}{٣٥}$

= ٦٠ + ٥,٧١ = ٦٥,٧١

حساب الوسيط بواسطة التكرار المتجمع الهابط:

جدول التكرار المتجمع الهابط

الحدود الدنيا	تكرار متجمع هابط
٤٠ فأكثر	١٠٠
٥٠ فأكثر	٩٥
٦٠ فأكثر	٧٠ ك اللاحق
٧٠ فأكثر	٣٥ ك السابق
٨٠ فأكثر	١٠

معادلة الوسيط فى حالة التكرار المتجمع الهابط هى:

الوسيط = نهاية الفئة الوسيطة −

$$\frac{\text{ترتيب الوسيط − التكرار المتجمع السابق}}{\text{التكرار الأصلى للفئة الوسيطة}} \times \text{مدى الفئة}$$

$$\text{الوسيط} = ٧٠ - \frac{٥٠ - ٣٥}{٣٥} \times ١٠$$

$$= ٧٠ - \frac{١٥٠}{٣٥} = ٧٠ - ٤,٢٩$$

$$= ٦٥,٧١$$

وهى نفس النتيجة التى حصلنا عليها باستخدام جدول التكرار المتجمع الصاعد

استخدام منحنى المتجمع الصاعد والهابط فى إيجاد قيمة الوسيط:

يمكن باستخدام القيم الحقيقية أو النسب المئوية لتكرار الفئات إيجاد الوسيط بالرسم، إلا أن هذه الطريقة تعتبر أقل طرق حساب الوسيط دقة. وبنفس طريقة رسم المنحنى المتجمع الصاعد والهابط والتى أشرنا إليها بإيجاز فى الفصل السابق. يقوم الدارس برسم المنحنين الصاعد والهابط معاً لإيجاد الوسيط حتى يمكن تحقيق قدر مناسب من الدقة. ومن نقطة التقاء المنحنيين يتم إسقاط عمود على المحور الأفقى (فئات) فيقطعه فى نقطة بعدها السينى يمثل قيمة الوسيط.

مثال (١٣):

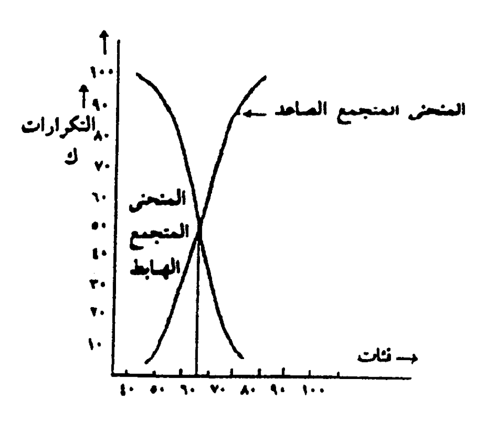
من بيانات جدول رقم (٤-١) فى المثال السابق أوجد الوسيط باستخدام منحنى المتجمع الصاعد والهابط.

الحـــل :

١- عمل جدول تكرار متجمع صاعد وهابط.

٢ - رسم المنحنى المتجمع الصاعد والهابط بالطريقة التى سبق شرحها فى الفصل الثالث. وكما هو واضح فى الشكل رقم (٤ -٦).

ف	ك	الحدود العليا	تكرار متجمع صاعد	الحدود الدنيا	تكرار متجمع هابط
٤٠ –	٥	أقل من ٥٠	٥	٤٠ فأكثر	١٠٠
٥٠ –	٢٥	أقل من ٦٠	٣٠	٥٠ فأكثر	٩٥
٦٠ –	٣٥	أقل من ٧٠	٦٥	٦٠ فأكثر	٧٠
٧٠ –	٢٥	أقل من ٨٠	٩٠	٧٠ فأكثر	٣٥
٨٠ – ٩٠	١٠	أقل من ٩٠	١٠٠	٨٠ فأكثر	١٠
مج	١٠٠				



شكل رقم (٤ – ٦) إيجاد الوسيط بالرسم

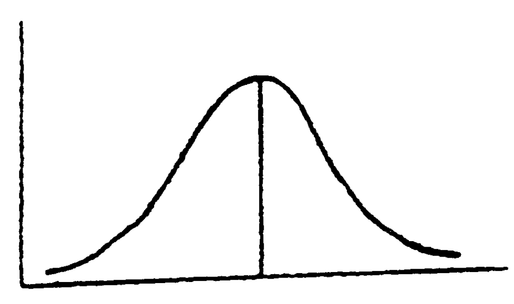
العلاقة بين المتوسطات الثلاثة للنزعة المركزية :

نستنتج من دراستنا للمقاييس الثلاثة للنزعة المركزية، إنها تنطبق جميعها أى تتساوى فى القيمة فى حالة المنحنيات المتماثلة مثل المنحنى الجرسى. ودائماً ما يقع الوسيط بنسب متفاوتة الشكل وفقاً لشكل التوزيعات بين المنوال والمتوسط الحسابى. أيضاً نجد أن المتوسط الحسابى ينحاز إلى طرف التوزيع أسرع من

الوسيط وتقدر المسافة بين المنوال والوسيط بحوالى ثلثى المسافة بين المنوال والمتوسط الحسابى. ومن ثم أمكن ببعض العمليات الرياضية البسيطة إيجاد معادلة تربط بين المقاييس الثلاثة والتى يمكن بواسطتها حساب أى منها إذا أمكن حساب المقياسين الآخرين حيث:

المنوال = الوسط الحسابى - ٣ (الوسط الحسابى - الوسيط)

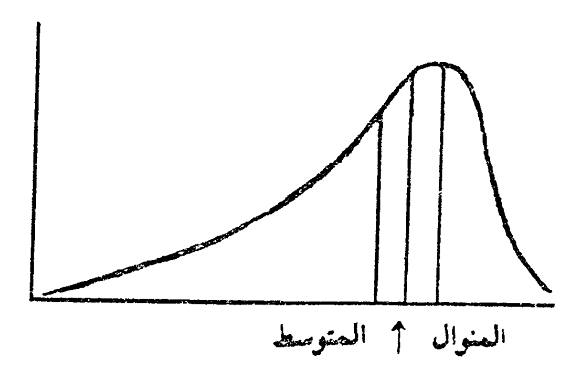
كذلك يمكن التعرف على شكل التوزيع التكرارى من حيث التماثل أو الإلتواء بالمقارنة بين قيم المتوسطات الثلاثة كما يمكن استنباط ذلك أيضاً من العلاقة السابقة . فكما قلنا يكون التوزيع متماثلا إذا تساوت قيم المنوال والوسيط والمتوسط الحسابى. والقيمة هنا هى المسافة على المحور السينى وعند النقطة التى يقطعها العمود الساقط من قمة المنحنى الجرسى على المحور السينى، كما يتضح ذلك من الشكل رقم (٤-٧) .



(المنوال ، الوسيط ، المتوسط)

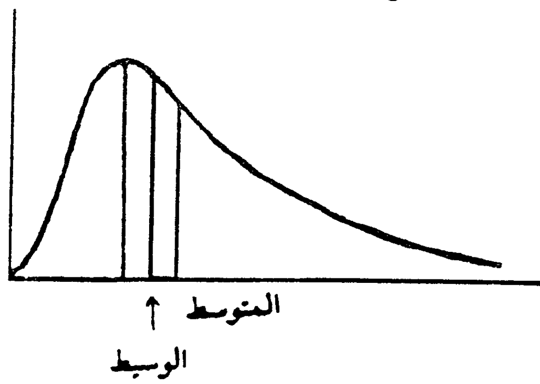
شكل رقم (٤ - ٧) منحنى متماثل

وعندما تكون قيمة المتوسط الحسابى أصغر من قيمة المنوال بفارق يقل عن الصفر فإن التوزيع يكون سالب الإلتواء Negative skewness كما يتضح ذلك من الشكل التالى رقم (٤-٨) .



شكل رقم (٤-٨) التواء ناحية اليسار (سالب)

أيضاً كلما كانت قيمة المتوسط الحسابى أكبر من قيمة المنوال بفارق موجب أى أكبر من الصفر فإن التوزيع التكرارى يكون موجب الإلتواءPositive skewness . كما يتضح ذلك من الشكل رقم (٤ -٩) .



شكل رقم (٤ ٩) التواء ناحية اليمين (موجب)

ملاحظات على مقاييس النزعة المركزية :

المقياس	المزايا	العيوب
المنوال	١ – يعتبر أفضل المقاييس بل المقياس الوحيد للمتوسط فى حالة البيانات الوصفية. ٢ – يعتبر أفضل المقاييس شيوعاً للتعبير عن شكل وتوزيع البيانات . ٣ – يمتاز بسهولة حسابه . ٤ – لايتأثر بالقيم الشاذة. ٥ – يمكن حسابه فى حالة التوزيعات المفتوحة خاصة البيانات غير الرقمية فيعتبر أفضل المتوسطات الثلاثة تمثيلاً لتلك البيانات .	١ – يتأثر بتغيير أطوال الفئات مما يقلل من أهميته ويحد من استخداماته . ٢ – مقياس غير مستقر تتوقف قيمته فى حالة التوزيعات التكرارية على طريقة التبويب .
الوسط الحسابى	١ – يعتبر أفضل المقاييس الثلاثة لتقدير الوسط الحسابى الحقيقى للمجتمع الأصلى .	١ – يصعب تقدير المتوسط الحسابى بدقة من التوزيعات التكرارية المفتوحة .

المقياس	المزايـــــا	العـــــيوب
الوسط الحسابى	٢ - يتميز عن الوسيط والمنوال بأنه يستخدم جميع البيانات المتاحة عن الظاهرة موضوع الدراسة . ٣ - يعتبر أفضل المقاييس الثلاثة فى حالة القياسات الكمية .	٢ - يتحيز أحياناً للقيم المتطرفة خاصة إذا كان حجم العينة صغير . ٣ - استبعاد كل الفئات المفتوحة فى حساب المتوسط .
الوسيط	١ - أفضل المقاييس الثلاثة إذا كانت القياسات التى تسجل عن الظاهرة ترتيبية . ٢ - يفضل عن المتوسط الحسابى كلما إزدادت درجة إلتواء التوزيعات التكرارية ، نظراً لأن الوسيط لايتأثر غالباً بالقيم المتطرفة للظاهرة . ٣ - يتميز بإمكانية استخدامه حتى فى حالة عِدم معرفة القيم الكبرى أو الصغرى فى جداول التوزيعات .	١ - صعوبة استخدامه فى عمليات جبرية . ٢ - لايمكن حساب الوسيط العام لعدة مجموعات من البيانات.

المتوسط المرجح Weighted mean

فى حالة العينات كبيرة الحجم حيث تكون البيانات منظمة إما فى جداول على شكل أرقام أو نسب مئوية . وتضم البيانات المبوبة مجموعات من البيانات والمطلوب حساب المتوسط الحسابى لها . وفى هذه الحالة يطلق على هذا المتوسط متوسط المجموعة (أو المتوسط الكبير) ونرمز له بالرمز (م) حتى نميزه عن المتوسط الحسابى البسيط لاى مجموعة (سَ). وكما نعلم أنه قد يسهل حساب المتوسط الحسابى لجميع المجموعات للبيانات المبوبة إذا كانت كل مجموعة منها متساوية مع الأخرى فى عدد مفرداتها . ففى هذه الحالة نحسب المتوسط الحسابى لكل جماعة . ثم نقوم بجمع المتوسطات لجميع الجماعات ثم نقسمها على عددها فنحصل على المتوسط الكبير . وهذه أبسط حالات حساب متوسط المجموعة .

من جهة أخرى إذا كان عدد مفردات مجموعة يختلف عن العدد فى المجموعة الأخرى ، فى هذه الحالة يتم حساب متوسط المجموعة (م) باتباع الخطوات التالية كما هو موضح فى المثال الآتى :

مثال (٤) :

يوضح الجدول رقم (٤ – ٣) توزيع عينة من الوحدات المعيشية The households تبعاً للحجم مقارنة بين عامى ١٩٧٨ ، ١٩٩٨ م . والمطلوب حساب متوسط المجموعة (م) ؛ ومنه فسر كيف يتغير متوسط حجم الوحدة المعيشية خلال الفترة مابين عامى ١٩٧٨ ، ١٩٩٨ ؟

جـدول رقـم (٤ - ٣)

حجم الوحدة المعيشية (س)	(ك) التكرارات ١٩٧٨ ك (٪)	١٩٩٨ ك (٪)
١	١٢	٢٥
٢	٣٠	٣٤
٣	٢٣	١٧
٤	١٩	١٦
٥	٩	٦
٦	٧	٢

بيانات هذا الجدول نسب مئوية

توضح بيانات هذا الجدول المدى الواسع فى توزيعاتها والتباين فيما بينها ابتداء من أسرة يمثلها فرد واحد من حيث الحجم حتى حجم أسرة يبدأ من ستة أفراد فأكثر دون تحديد لأقصى حجم مما يصعب معه حساب المتوسط الحسابى البسيط (سَ) لأن قيمته لاتكون حقيقية . ومن هنا تكون ميزة استخدام المتوسط الكبير (م) لانه يكون أكثر دقة مع البيانات المبوبة التى تتباين بداخلها الأعداد أو النسب المئوية بين مجموعة وأخرى كما هو الحال فى تباين حجم الوحدات المعيشية فى هذا المثال .

يتضح الاختلاف أيضاً فى حجم الوحدة المعيشية اذا ماقورنت النسب المئوية لعامى ١٩٧٨ ، ١٩٩٨ . فبالنسبة لحجم الوحدة المعيشية التى تضم فرداً يعيش بمفرده تصل النسبة المئوية لهذا الحجم فى عام ١٩٩٨ إلى الضعف تقريباً من نسبتها المئوية فى عام ١٩٧٨ (٢٥٪ فى عام ١٩٩٨ مقابل ١٢٪ عام ١٩٧٨) . من جهة أخرى تنخفض النسبة المئوية لحجم الوحدة المعيشية كبيرة الحجم (+٦)

من ٧٪ فى عام ١٩٧٨؛ إلى ٢٪ فى عام ١٩٩٨ . **من ذلك يتضح وجود تحول واضح ذا دلالة** نحو الحجم الأصغر للوحدة المعيشية **وهذا ينعكس بدوره على متوسط حجمها** . كما ستكشف عنه الخطوات الآتية فى **حساب متوسط الجماعة فى هذا المثال** الذى يتصف بالتباين الواضح فى التكرار النسبى للبيانات .

خطوات حساب متوسط المجموعة :

١ – اضرب حجم الوحدة المعيشية (س) **فى وزنها أو تكراراتها المناظرة . فى هذا المثال** تمثل التكرارات (ك) النسبة المئوية لحجم الوحدات المعيشية **(لعام ١٩٧٨) .**

(١٢×١) + (٣٠×٢) + (٢٣×٣) + (١٩×٤) + (٩×٥) + (٧×٦) = ٣٠٤

٢ – أقسم الناتج من الخطوة الأولى على مجموع التكرارات أو الاوزان

(١٢ + ٣٠ + ٢٣ + ١٩ + ٩ + ٧) تحصل على متوسط المجموعة م .

م = $\frac{٣٠٤}{١٠٠}$ = ٣,٠٤ فرد لكل وحدة معيشية

٣ – كرر الخطوات السابقة فى حساب متوسط المجموعة لبيانات عام ١٩٩٨ وسوف تحصل على متوسط مجموعة يساوى ٢,٥٠ فرد لكل وحدة معيشية . ويفسر هذا الفرق بين المتوسطين اتجاه حجم الوحدة المعيشية إلى الانخفاض النسبى فى عام ١٩٩٨ مقارنة بعام ١٩٧٩ .

متوسط المجموعة (سَ) = $\frac{\text{مجـ س ك}}{\text{مجـ ك}}$

الحـــل :

س	١٩٧٨		١٩٩٨	
	ك	س ك	ك	س ك
١	١٢	١٢	٢٥	٢٥
٢	٣٠	٦٠	٣٤	٦٨
٣	٢٣	٦٩	١٧	٥١
٤	١٩	٧٦	١٦	٦٤
٥	٩	٤٥	٦	٣٠
٦	٧	٤٢	٢	١٢
مجـ	١٠٠	٣٠٤	١٠٠	٢٥٠

متوسط حجم الوحدة المعيشية فى عام ١٩٧٨ = $\frac{\text{مجـ س ك}}{\text{مجـ ك}}$

= $\frac{٣٠٤}{١٠٠}$

= ٣,٠٤ فرداً لكل أسرة

متوسط حجم الوحدة المعيشية فى عام ١٩٩٨ = $\frac{٢٥٠}{١٠٠}$ = ٢,٥ فردا لكل أسرة .

أما إذا كانت البيانات المجدولة مفتوحة النهاية كما فى الجدول رقم (٤ - ٤) فهذا يعنى أن حجم الوحدة المعيشية قد يكون ستة أفراد فأكثر (٦+) أى ٦ ، ٧ ، ٨، ٩ ، ...... الخ . ولحساب المتوسط المرجح فى هذه الحالة يختار رقماً أكبر من (٦) وليكن رقم (٩) من ثم سوف يتغير مجموع (س ك) وبالتالى قيمة الوسط الحسابى المرجح .

المتوسط المرجح Weighted Mean

جدول رقم (٤ – ٤) التوزيع التكرارى للوحدات المعيشية
وفقاً للحجم فى عامى ١٩٧٨ ، ١٩٩٨

س	١٩٧٨		١٩٩٨	
	ك	س ك	ك	س ك
١	١٢	١٢	٢٥	٢٥
٢	٣٠	٦٠	٣٤	٦٨
٣	٢٣	٦٩	١٧	٥١
٤	١٩	٧٦	١٦	٦٤
٥	٩	٤٥	٦	٣٠
٦+ افتراضيا=٩	٧	٦٣	٢	١٨
مجـ	١٠٠	٣٢٥	١٠٠	٢٥٦

∴ المتوسط المرجح لحجم الأسر فى عام ١٩٧٨ = $\frac{٣٢٥}{١٠٠}$

= ٣,٢٥ أفراد لكل أسرة

المتوسط المرجح لحجم الأسر فى عام ١٩٩٨ = $\frac{٢٥٦}{١٠٠}$

= ٢,٥٦ أفراد لكل أسرة

متوسط الجماعات المرتبطة The Mean of Combined Groups

يستخدم متوسط الجماعة فى البيانات المبوبة التى تشتمل على جماعات مرتبطة تشترك فى ظاهرة معينة رغم الإختلاف بينها من حيث عدد المفردات والنوعية . ويوضح المثال التالى كيف يمكن حساب متوسط المجموعة للجماعات المرتبطة .

مثال (١٥) :

فى امتحان مادة الانثروبولوجيا الحضرية ، كان عدد الطالبات ١٠٨ طالبة ، وعدد الطلاب ٤٧ طالباً تقدموا لهذا الامتحان . وحصلت الطالبات على درجات فى هذه المادة متوسطها الحسابى $\bar{س}_{أ}$ = ٤٥,٢٦ فى حين بلغ المتوسط الحسابى لدرجات الطلاب ، $\bar{س}_{ذ}$ = ٥٤,٨٩ . والمطلوب حساب المتوسط العام للطالبات والطلاب داخل الجماعة (م) (أو متوسط التوزيع الكلى للطلاب وا لطالبات) ؟

الحـــل :

$$\bar{م} = \frac{(\text{مجـ ك}_{ذ}\ \bar{س}_{ذ}) + (\text{مجـ ك}_{أ}\ \bar{س}_{أ})}{\text{مجـ ك}_{ذ} + \text{مجـ ك}_{أ}}$$

حيث أن : مجـ ك $_{أ}$ = مجموع تكرارات الاناث

مجـ ك $_{ذ}$ = مجموع تكرارات الذكور

$\bar{س}_{أ}$ = المتوسط الحسابى للاناث

$\bar{س}_{ذ}$ = المتوسط الحسابى للذكور

$$\therefore \text{المتوسط} = \frac{(٤٧)\ (٥٤,٨٩) + (١٠٨)\ (٤٥,٢٦)}{٤٧ + ١٠٨}$$

= ٤٨,١٨ درجة .

جدول رقم (٤ - ٥)
التوزيع التكرارى لدرجات الامتحان النهائى لمادة الانثروبولوجيا الحضرية

الدرجة (س)	ك
٧٩	٢
٦٨	١
٦٥	٥
٦٣	١٢
٦٠	١٧
٥٨	١١
٥٣	٢٠
٥٢	١٣
٥١	٢٣
٤٧	١٨
٤٥	١٤
٤٠	١١
٣٨	٥
٣٥	٣
	١٥٥

المنوال mode :

المنوال لمجموعة من القيم هو القيمة التى تقابل أكبر تكرار فى المجموعة أو بمعنى آخر هى القيمة الأكثر شيوعاً.

المتوسط الحسابى Arithmetic Mean :

هو أكثر مقاييس النزعة المركزية استخداماً ونحصل عليه بقسمة مجموع القيم على عددها.

الوسيط Median :

تعنى كلمة الوسيط منتصف الشىء. فهو القيمة التى تقع فى المنتصف تماماً بعد ترتيب القيم تنازلياً أو تصاعدياً. أى القيمة التى يكون عدد القيم الأخرى الأقل منها مساوياً تماماً لعدد القيم الأعلى منها. والوسيط أفضل مقاييس النزعة المركزية استخداماً فى حالة التواء التوزيع.

تمــــــارين

١ - فيما يلى جدول توزيع تكرارى (٥٠٠) عاملاً بأحد المصانع والمطلوب قياس متوسط العمر لهم بثلاث طرق مختلفة. مع بيان افضل المتوسطات استخداماً فى تمثيل أعمارهم وسبب هذا التفضيل.

الأعمـــــار	عدد العــــمال
١٥ -	٤٥
٢٥ -	٢٣٥
٣٥ -	١٥٠
٤٥ -	٥٠
٥٥ - ٦٥	٢٠
	٥٠٠

٢- احسب كل من الوسيط والمتوسط الحسابى والمنوال للأجور التالية. ثم بين أى المقاييس أفضل.

٣,١٠	٢,٥٧	٢,٥٠	٧,٨٠	٣,٧٥
٣,٢٨	٣,٩٦	٢,٥٧	٣,٦٠	٢,٥٧

٣ – الجدول الآتى يبين عدد المدمنين حسب فئاتهم العمرية فى إحدى المدن عام ١٩٩٨ والمطلوب :

الفئات	عدد الحالات
١٥–	٧٠
٢٠–	١٢٠
٢٥–	٤٦٠
٣٠–	٢٠٠
٣٥–	١٦٧
٤٠–	٨٢
٤٥–	٢٧
٥٠–٥٥	١٥

(أ) عمل جدول تكرارى نسبى ومئوى.

(ب) رسم المدرج التكرارى لهذا التوزيع وأوجد منه قيمة المنوال ثم حقق الناتج حسابياً.

٤– فيما يلى توزيع تكرارى لبيانات افتراضية عن الدخل السنوى لعينة من الأسر المنوال:

فئات الدخل بالجنيه	ك
٧٠–	١٦
٩٠–	١١٨
١١٠–	٤٠٦
١٣٠–	١٢٦٤
١٥٠–	١٦١١

فئات الدخل بالجنيه	ك
١٧٠-	١١٥١
١٩٠-	٤٩٧
٢١٠-	٢٣٨
٢٣٠-	٧٠
٢٥٠-٢٧٠	٣٢

٥- فيما يلى درجات خمسين طالباً فى الامتحان النهائى لمادة الإحصاء الاجتماعى، والمطلوب جدولة هذه البيانات فى توزيع تكرارى مئوى ثم احسب المتوسط الحسابى والمنوال من هذا الجدول:

١٤	١٥	١٦	١٩	١٧
١٧	١٨	١٤	١٤	٢٠
١٢	١٤	١٢	١٩	١٤
١٠	١٢	١٥	٢٠	٨
١٣	١١	١٩	١٥	٩
٨	١٢	٢٠	١٦	١٠
١٤	١٤	١٧	١٧	٢٠
١٠	١١	٢٠	١٣	١٥
١٥	١٥	١٠	٨	١٣
١٢	١٢	١٣	٩	١٧

٦- فيما يلى مجموعة من القيم والمطلوب حساب المتوسطات الثلاثة:

٦، ٢، ٤ ، ٧ ، ٢٠ ، ٤ ، ٧، ٨، ٨ ، ٩، ١٢، ٧، ٥، ٦، ١٩ ، ٢٠، ٦ ، ٨ ، ١١، ٧، ٩.

٧ - فيما يلى درجات ٣٠ طالباً فى امتحان مادة مناهج البحث:

٦٥	٧٤	٣٥	٣٨	٧٠
٩٥	٥٨	٧٤	٧٠	٦٥
٧٣	٣٥	٧٠	٧٠	٦٧
٤٦	٧٠	٤٨	٣٥	٦٦
٣٠	٩٦	٣٥	٩٦	٤٨
٩٥	٤٨	٩٥	٧٠	٩٦

المطلوب جدولة هذه الدرجات الخام وحساب الوسيط والمتوسط الحسابى. على افتراض أن المتغير متصل.

٨ - تمثل البيانات الآتية أوزان افتراضية لعينة من المبحوثين والطلوب جدولة تلك الأوزان وحساب مقاييس النزعة المركزية. مع توضيح العلاقة بين هذه المقاييس بالرسم. ومناقشة تلك العلاقة.

٨٠	٨٥	٨٤	٦٠	٦٥
٤٠	٣٢	٨١	٦٠	٥٣
٩٥	٩٥	٨٠	٩٥	٨٤
٦٣	٨٠	٧٣	٨٠	٩١
٧٠	٥٢	٦٣	٤٠	٦٣

٩- احسب المتوسط الحسابى والوسيط من الجدول التكرارى الآتى مع توضيح أنسب مقاييس النزعة المركزية فى حالة وجود التواء فى توزيع القيم ، ثم وضح التجمع الصاعد والهابط بالرسم:

الفئات	ك
١٠-	٢
٢٠-	١٩
٣٠-	٩
٤٠-	٣
٥٠-	٢١
٦٠-	٧
٧٠-	٨
٨٠-	١٥
٩٠-١٠٠	٦

١٠- احسب المتوسط الحسابى والمتوسط الفرضى والمنوال من بيانات جدول التوزيع التكرارى التالى:

الفئات	ك
٢٠-	١
٢٥-	٢
٣٠-	٥
٣٥-	٢٠
٤٠-	٢٢
٤٥-	٤٢
٥٠-	٣٠

الفئات	ك
٥٥-	١٠
٦٠-	١٥
٦٥-٧٠	٦

١١- أكمل ما يأتى :

(أ) تشير عبارة النزعة المركزية إلى .... حالة فى وجود التوزيع .

(ب) المقاييس الثلاثة الأساسية للنزعة المركزية هى المنوال ، .... ، ....

(جـ) أن الغرض من جميع مقاييس النزعة المركزية هو .... للتوزيع الكلى للقيم بواسطة وصف أكثر القيم المتطابقة داخله .

(د) على النقيض من المنوال ، فإن .... يمثل دائماً المركز الحقيقى لتوزيع القيم .

(هـ) أن .... يعتبر أكثر المقاييس استخداماً للنزعة المركزية إلا أنه يتطلب مراجعة كاملة فقط فى حالة استخدامه مع مستوى بيانات ....

(و) يتأثر مقياس .... بكل قيمة داخل التوزيع .

(ز) عندما يكون للتوزيع قيما قليلة عالية القيمة بشكل شاذ عن باقى القيم ، يطلق على هذا التوزيع بأنه .... الالتواء وأن .... سوف يكون له أكبر القيم عن ....

١٢ – أمامك عدة اختيارات احداها يمثل الاجابة الصحيحة على السؤال المطلوب الاجابة عليه بوضع دائرة على الاختيار الصحيح فيمايلى :

– ان السبب الرئيسى لحساب مقاييس النزعة المركزية يتمثل في :

(أ) تلخيص المتغيرات الفردية .

(ب) ايجاد قيمة متوسطة .

(جـ) معرفة خصائص المتغير .

(د) لا إجابة من الاجابات الثلاثة السابقة .

(هـ) جميع الإجابات الثلاث السابقة .

– يعرف الوسيط بأنه النقطة التى :

(أ) تمثل أعلى التكرارات المشاهدة .

(ب) تمثل القيمة التى عندما يتم طرحها من المتوسط يكون المنوال هو النتيجة .

(جـ) تمثل المحل المركزى داخل التوزيع .

(د) تتمثل فى جميع الإجابات الثلاث السابقة .

(هـ) لا تتمثل فى جميع الإجابات السابقة .

– إن أكثر المتغيرات استخداماً فى البحث الاجتماعى هو الرقم المتوسط لسنوات التعليم التى يتلقاها المصريون فى مراحل التعليم المختلفة . وهذا الرقم هو :

(أ) المنـــوال .

(ب) الوســـيط .

(جـ) المتوســط .

(د) ليس واحداً من المقاييس السابقة .

(هـ) واحد فقط من المقاييس السابقة .

– لو أن أستاذ مادة النظرية الاجتماعية قد أعلن أن كل درجة حصلت عليها طالبات السنة الثالثة من قسم الاجتماع قد زادت سبع درجات لأنه استبعد سؤالين من أسئلة الامتحان واضيفت درجاتهما إلى باقى أسئلة الامتحان ففي هذه الحالة ماذا حدث للمتوسط الجديد ؟

(أ) لم يتغير .

(ب) يساوى المتوسط القديم مع إضافة قيمة تساوى ($\frac{٧}{ن}$)

(جـ) يساوى المتوسط القديم مضافاً إليه (٧) .

(د) يساوى المتوسط القديم مضافاً إليه قيمة مقدارها $\frac{١٤}{٢}$ درجة .

(هـ) لا توجد معلومات كافية للاجابة على السؤال .

– فى إحدى التوزيعات النوعية ، كانت قيمة المنوال (٧٥) ، والوسيط (٧٠) ، والمتوسط (٦٥) . فهل يكون هذا التوزيع :

(أ) اعتيادياً Normal .

(ب) يتصف بالالتواء الموجب .

(جـ) يتصف بالالتواء السالب .

(د) متماثل Symmetrical .

(هـ) لايتصف بأى صفة من الصفات الأربع السابقة .

- يستخدم المنوال احصائياً فى قياس النزعة المركزية عندما تكون القياسات المعطاة

(أ) اسمية nominal .

(ب) ترتيبية .

(جـ) فلوية .

(د) نســبة .

(هـ) جميع الخصائص السابقة.

- لو كان التوزيع المعطى متصفاً بالتماثل ، فإن أفضل مقاييس النزعة المركزية استخداماً فى هذه الحالة هو :

(أ) المنـــوال .

(ب) الوســـيط .

(جـ) المتوســط .

(د) جميع المقاييس المذكورة سابقاً .

- أحسب المنوال من الحسابات الآتية:

١ - المتوسط الحسابى = ٤٦ الوسيط = ٤٠

٢ - المتوسط الحسابى = ٣٥ الوسيط = ٣٢

٣ - المتوسط الحسابى = ٢١٠ الوسيط = ٢٠٥

* * * * *

الفصل الخامس
مقاييس التشتت



الفصل الخامس

مقاييس التشتت

مقدمة

تناولنا فى الفصلين الثالث والرابع خاصيتين أساسيتين من خصائص التوزيع هما الشكل والتعبير عنه بالرسومات البيانية ، ثم مقاييس النزعة المركزية . من ثم تتبقى خاصية ثالثة هى التشتت أو درجة تباين القيم داخل التوزيع محل الدراسة. فاذا كانت مقاييس النزعة المركزية تمد الباحث بقيمة واحد تصف حالة التوزيع للقيم جملة واحدة ، فان للتشتت أهمية فى قياس الفروق الفردية داخل التوزيع، بمعنى آخر توضح مقاييس التشتت (التباين) Measures of Dispersion (variation) درجة التقارب أو التباعد بين القيم فى العينة موضوع الدراسة . وتوجد مقاييس عديدة متاحة فى الإحصاء الوصف لقياسس التباين للمتغير المتصل هى:

١ – المدى (المدى المطلق) The Range

٢ – الانحراف الربيعى The quartile Deviation

٣ – الانحراف المتوسط The mean deviation

٤– التباين The Variance

تمثل مقاييس التشتت مؤشرات دالة على وقوع الاختلافات لتوزيع ظاهرة ما محل الدراسة . وجدير بالذكر أن مقاييس النزعة المركزية والأشكال البيانية لاتكفى لوصف توزيع ما أو لعقد مقارنات بين مجموعة وآخرى. وفيما يلى مثالاً يوضح أنه رغم تساوى قيم المتوسط الحسابى لمجموعتين فهذا لايعنى وجود اتساق فى القيم حول المتوسط، حيث أن واقع التوزيع يشير إلى اختلاف بينهما من حيث تشتت المفردات.

مـثال (١) :

أجرى باحث اجتماعى دراسة حول التماسك الأسرى لعدد من الإطباء والمدرسين وكانت النتائج على النحو التالى:

الإطباء	المدرسون
١٣	٤
١٥	٣٧
١٣	٢٩
١٤	٧
١٧	١٠
١٤	٣٤
١٦	٥
١٨	٩
١٩	٨
١١	٧
مجـ ١٥٠	١٥٠

$$\bar{س}_{الإطباء} = \frac{١٥٠}{١٠}$$

$$\bar{س}_{المدرسين} = \frac{١٥٠}{١٠}$$

يتضح تساوى قيمة الوسط الحسابى للمجموعتين. إذا تفحصنا توزيع الدرجات حول وسطها الحسابى أى مدى قربها أو بعدها منها لوجدنا توزيع القيم اكثر تشتتا فى مجموعة المدرسين عنه فى مجموعة الاطباء. فإذا استخدمنا الوسط الحسابى فقط لعقد المقارنة بين المجموعتين لكان مضللاً . ومن ثم كان ضرورياً الاهتمام بالتباين أوالتشتت فى التوزيعات محل الدراسة .

وسوف نتناول فى هذا الفصل مقاييس التشتت وخصائص كل منها ونقاط القوة والضعف فى كل مقياس. وينقسم الفصل إلى :

١ - مقاييس التباين للمتغير المتصل

٢ - مقياس التباين للمتغير المتقطع

٣ - المفاهيم الأساسية

٤ - تمارين

مقاييس التباين للمتغير المتصل :

المـــــدى :

يعتبر المدى أبسط مقاييس التشتت وأسهلها فى الحسابات وله مميزات كما أن له عيوباً. قد يكون مقياس المدى نافعا فى الحالات التى تتطلب سرعة فى الحسابات والحصول على مؤشرات أولية عن التشتت لتوزيع ما. كما يكون مفيدا للمبتدئين والذين لا تتوفر لديهم المهارة الإحصائية الكافية لاستخدام مقاييس تشتت اكثر تعقيداً.

يكون استخدام مقياس المدى أكثر نفعا فى حالة البيانات التى لا تتطلب معالجة إحصائية متقدمة أو البيانات التى سبق دراسة خصائصها وتتطلب مجرد الإلمام بما تتصف به من تشتت.

كيفية استخدام المدى فى قياس التشتت:

يعرف المدى بأنه الفرق الحسابى المطلق (أى بدون استخدام الإشارات الموجبة أو السالبة للقيم) بين أعلى قيمة وأقل قيمة فى التوزيع فى حالة البيانات غير المبوبة. أو الفرق بين الحد الأعلى للفئة العليا والحد الأدنى للفئة الدنيا فى التوزيعات التكرارية . كما يعرف بأنه الفرق بين مركز الفئة الاعلى ومركز الفئة الأدنى.

مثال (٢) :

فيما يلى أوراق لعددمن الاطفال والمطلوب حساب المدى المطلق

١٣ ، ١٥ ، ٣٥ ، ١٧ ،

٩ ، ٢٠ ، ١٨ ، ١٦

خطوات الحـل:

١ - ترتيب القيم إما تصاعديا أو تنازليا

٢ - تحديد أكبر قيمة وأصغر قيمة فى التوزيع

٣ - طرح أصغر قيمة من أكبر قيمة

ترتيب القيم

٩ ١٣ ١٥ ١٦ ١٧ ١٨ ٢٠ ٣٥

∴ المدى المطلق = ٣٥ - ٩ = ٢٤ كيلو جرام.

من عيوب مقياس المدى أنه يعتمد فى حساب التشتت على أعلى قيمة وأصغر قيمة فى التوزيع. وقد يصعب الحصول عليهما فى العينة البحثية . ومن ثم تلعب الصدفة دوراً فى تكرار حدوث أعلى وأقل القيم داخل توزيع بعينة . فإذا افترضنا أن باحثاً أراد أن يعرف التشتت فى توزيع الثروة داخل مجتمع محلى ، ولا يوجد بداخله سوى مليونير واحد فقط . فلو اختار الباحث عينة من عشرة أو عشرين فرداً من هذا المجتمع، فإن احتمالات أن تشمل هذه العينة المليونير الوحيد ستكون ضعيفة جداً . من ثم تكون قيمة المدى الدالة على التشتت فى توزيع الثروة مضللة (Graham, 1994: 80 & Blalock 1972 : 78) من عيوب المدى أيضا ، أن الباحث عندما يستخدمه لا يستطيع أن يتعرف على درجة التباين للقيم الواقعة بين أعلى القيم وأدناها فى التوزيع . لهذا السبب نجد أن الباحثين فى استخدامهم للمدى يضيفون إلى الفرق بين أعلى قيمة وأقل قيمة مقدار الواحد الصحيح حتى تغطى قيمة المدى الدالة على تشتت التوزيع أعل القيم وأقل القيم. ويتم حساب المدى باستخدام المعادلة الآتية:

المدى = (أعلى قيمة - أقل قيمة) + ١

مثال (٣) :

وضح باستخدام مقياس المدى أى التوزيعين الآتيين أكثر تشتتا من الآخر:

التوزيع الأول	١١	١٦	١٨	٢٣	٢٩	٣١	٣٧
التوزيع الثانى	١٨	١٩	٢١	٢٣	٢٤	٢٦	٢٩

الحــــل :

المدى للتوزيع الأول = (٣٧ - ١١) + ١ = ٢٧

المدى للتوزيع الثانى = (٢٩ - ١٨) + ١ = ١٢

أى أن التوزيع الأول اكثر تشتتا فى قيمه من التوزيع الثانى. (Hinkle et al., 42 : 1979) علما بأن التوزيعين لهما قيمة متماثلة للوسيط.

من عيوب استخدام المدى عدم صلاحيته للتطبيق على المجتمع الأصلى والعينات كبيره الحجم بينما يسهل حسابه على العينات صغيرة الحجم حيث تكون الفرصة أفضل لاشتمالها على قيم شاذة عليا ودنيا. من ثم نادرا مايستخدم علماء الاجتماع هذا المقياس الإتى الدراسات الكشفية أو الاستطلاعية فى معظم الاحيان. . Hinkle et al 1979 : 43 & Kurtz , 1983 71 - 72

الانحراف الربيعى:

نظرا لأن المدى يقوم على القيم الأعلى والأدنى من التوزيع، فضلاً عن إعتماده على عدد الحالات، فإنه يعتبر مقياساً غير مستقر . وهذا القصور فى المدى يمكن التغلب عليه بايجاد الانحراف الربيعى. ويعرف بنصف المسافة بين الربيعين الاول والثالث . فلو رمزنا للانحراف الربيعى بالرمز (ر) وللربيع الأول بالرمز ($ر_١$) ، وللربيع الثالث ($ر_٣$) يمكن حساب الانحراف الربيعى من المعادلة التالية:

$$\text{الانحراف الربيعى} = \frac{ر_٣ - ر_١}{٢}$$

ويستخدم مقياس الانحراف الربيعى عادة فى المستوى الترتيبى للبيانات . وهذا المقياس اكثر استخداما فى البحوث التربوية والنفسية ويندر استخدامه فى البحوث الاجتماعية (Graham, 1994 : 80, 81) وفيما يلى خطوات حساب الانحراف الربيعى.

١ – حساب الانحراف الربيعى من البيانات غير المبوبة:

مثال (٤) :

فيما يلى درجات عدد من الطالبات فى امتحان الإحصاء الاجتماعى والمطلوب حساب درجة التشتت باستخدام مقياس الانحراف الربيعى.

١١   ٩   ١٥   ١٧   ٨   ٧

١٤   ١٣   ٢٠   ١٨   ١٦   ٥

خطوات الحل :

١ – ترتيب القيم تصاعديا أو تنازليا

٢ – تحديد قيمة الربيع الادنى ($ر_{١}$)

٣ – تحديد قيمة الربيع الاعلى ($ر_{٣}$)

٤ – تطبيق المعادلة $\frac{ر_{٣} - ر_{١}}{٢}$

ترتيب القيم

٥   ٧   ٨   ٩   ١١   ١٣   ١٤   ١٥   ١٦   ١٧   ١٨   ٢٠

$$\text{قيمة } ر_{١} = \frac{\text{عدد القيم}}{٤} = \frac{ن}{٤}$$

$$= \frac{١٢}{٤} = ٣$$

∴ قيمة $ر_٣$ = القيمة الثالثة وهى حسب الترتيب ٨

قيمة $ر_٣ = \frac{ن}{٤} \times ٣$

$= \frac{١٢}{٤} \times ٣$

∴ قيمة $ر_٣$ هى القيمة التاسعة وقيمتها ١٦

الانحراف الربيعى $= \frac{ر_٣ - ر_١}{٢}$

$= \frac{١٦ - ٨}{٢} = ٤$

٢ – حساب الانحراف الربيعى من الجداول التكرارية

مثال (٥) :

يتضمن الجدول التالى توزيع تكرارى لاوزان عينة من الاطفال والمطلوب حساب قيمة الانحراف الربيعى:

الوزن	٤-	٨-	١٢-	١٦-	٢٠-	٢٤-	٢٨-٣٢
العدد	٣	٥	٦	٤	١٠	٧	٥

خطوات الحل:

١ – عمل جدول متجمع صاعد للبيانات.

٢ – تحديد موضع الربيع الأدنى $= \frac{مجـ ك}{٤}$

٣ – تحديد موقع الربيع الأعلى $= \frac{مجـ ك}{٤} \times ٣$

ف	ك	الحدود العليا	تكرار متجمع صاعد	
٤ –	٣	اقل من ٨	٣	
٨ –	٥	اقل من ١٢	٨	
١٢ –	٦	اقل من ١٦	١٤	موقع $ر_{١}$
١٦ –	٤	اقل من ٢٠	١٨	
٢٠ –	١٠	اقل من ٢٤	٢٨	
٢٤ –	٧	اقل من ٢٨	٣٥	موقع $ر_{٣}$
٢٨ – ٣٢	٥	اقل من ٣٢	٤٠	
مجـ	٤٠			

موقع الربيع الأدنى ($ر_{١}$) = $\frac{٤٠}{٤}$ = ١٠

$$\text{قيمة } ر_{١} = ح_{د} + \frac{\text{موقع } ر_{١} - \text{ك سابق}}{\text{ك الفئة الأصلية } ر_{١}} \times \text{طول الفئة}$$

$$= ١٢ + \frac{١٠ - ٨}{٦} \times ٤$$

= ١٢ + ٠,٣٣ (٤)

= ١٢ + ١,٣٣ = ١٣,٣ كيلوجرام

$$\text{قيمة } ر_{٣} = ح_{د} + \frac{\text{موقع } ر_{٣} - \text{ك سابق}}{\text{ك الفئة الأصلية } ر_{٣}} \times \text{طول الفئة}$$

موقع $ر_{٣}$ = $\frac{٤٠}{٤}$ × ٣ = ٣٠

$$\text{قيمة ر}_3 = ٢٤ + \frac{٣٠ - ٢٨}{٧} \times ٤$$

$= ٢٤ + ٠,٢٩ (٤)$

$= ٢٤ + ١,١٤ = ٢٥,١٤$ كيلوجرام

$$\therefore \text{ الانحراف الربيعى} = \frac{٢٥,١٤ - ١٣,٣}{٢} = ٥,٩٢ \text{ كيلوجرام}$$

الانحراف المتوسط:

يستخدم مقياس الانحراف المتوسط لقياس التشتت لأنه يتفادى أوجه القصور فى المقاييس السابقة، حيث يستخدم الانحراف المتوسط جميع القيم أى أنه يهتم بالانحرافات لكل قيمة عن وسطها الحسابى.

مثال (٦):

فيما يلى الدخل الشهرى (بمئات الجنيهات) لعدد من المشتغلين فى أحدى الشركات الاستثمارية والمطلوب حساب الانحراف المتوسط.

١٣	٩	١٥	١٧	٧	٨
١٤	٢٠	١٨	١٦	٥	٢

خطوات الحل :

١ – أحسب قيمة الوسط الحسابى للقيم جميعها

٢ – أطرح كل قيمة من الوسط الحسابى، كقيم مطلقة بمعنى إهمال الإشارة الجبرية لها (+ ، -)

٣ – أقسم اجمالى ناتج الطرح على عدد الحالات، تحصل على الانحراف المتوسط الدال على التباين فى توزيع البيانات.

حل المثال :

الدخل (س)	س َ – س
١٣	١
٩	-٣
١٥	٣
١٧	٥
٧	-٥
٨	-٤
١٤	٢
٢٠	٨
١٨	٦
١٦	٤
٥	-٧
٢	-١٠
مجـ ١٤٤	(٥٨) مع إهمال الإشارات الجبرية

$$\text{س َ} = \frac{\text{مجـ س}}{\text{ن}}$$

$$= \frac{١٤٤}{١٢} = ١٢$$

$$\text{الانحراف المتوسط} = \frac{\text{مجـ } / \text{س َ} - \text{س} /}{\text{ن}}$$

$$= \frac{٥٨}{١٢} = ٧ \text{ جنيهات تقريباً}$$

التباين والانحراف المعيارى:

يعرف التباين بأنه متوسط مربع انحرافات القيم عن وسطها الحسابى.
وسوف نرمز للتباين بالرمز (ع٢) بينما رمزه اللاتينى (σ^2) (سيجما)

١ - حساب التباين للبيانات غير المبوبة.

ونرمز للتباين بالرمز (ع٢) والانحراف المعيارى (ع).

سوف يتم حساب التباين باستخدام طريقتان: الأولى باستخدام المتوسط الحسابى للقيم، والثانية هى الطريقة المباشرة.

(أ) حساب التباين باستخدام المتوسط الحسابى للقيم فى المثال رقم (٦)

س	س - سَ	(س - سَ)٢
١٣	١	١
٩	٣-	٩
١٥	٣	٩
١٧	٥	٢٥
٧	٥-	٢٥
٨	٤-	١٦
١٤	٢	٤
٢٠	٨	٦٤
١٨	٦	٣٦
١٦	٤	١٦
٥	٧-	٤٩
٢	١٠-	١٠٠
مجـ ١٤٤	صفر	٣٥٤

$$\text{التباين} = \frac{\text{مجـ}(\bar{س} - س)^{٢}}{ن}$$

$$= \frac{٣٥٤}{١٢} = ٢٩,٥$$

الانحراف المعيارى (ع) عبارة عن الجذر التربيعى للتباين .

$$\therefore \quad ع = \sqrt{\text{التباين}}$$

$$ع = \sqrt{\frac{\text{مجـ}(\bar{س} - س)^{٢}}{ن}}$$

$$ع = \sqrt{٢٩,٥} = ٥,٤٣$$

(ب) حساب التباين والانحراف المعيارى للقيم غير المبوبة باستخدام الطريقة المباشرة

س	$س^{٢}$
١٣	١٦٩
٩	٨١
١٥	٢٢٥
١٧	٢٨٩
٧	٤٩
٨	٦٤
١٤	١٩٦
٢٠	٤٠٠
١٨	٣٢٤
١٦	٢٥٦
٥	٢٥
٢	٤
١٤٤	٢٠٨٢

$$\text{التباين} = \frac{1}{\text{ن}}\left[\text{مجـ س}^{2} - \frac{(\text{مجـ س})^{2}}{\text{ن}}\right]$$

$$= \frac{1}{12}\left[2082 - \frac{(144)^{2}}{12}\right]$$

$$= \frac{1}{12}\left(2082 - \frac{20736}{12}\right)$$

$$= \frac{1}{12}\left(2082 - 1728\right)$$

$$\therefore \text{ع}^{2} = \frac{354}{12} = 29{,}5$$

$$\text{ع} = \sqrt{29{,}5} = 5{,}43$$

حساب التباين والانحراف المعيارى للبيانات المبوبة:

وسوف نستخدم المعادلة التالية لحساب التباين والانحراف المعيارى .

$$\text{التباين} = \frac{\text{مجـ س}^{2}\text{ ك}}{\text{مجـ ك}} - \text{س}^{3}$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ س}^{2}\text{ ك}}{\text{مجـ ك}} - \bar{\text{س}}^{2}}$$

وباستخدام بيانات المثال رقم (٥) أحسب قيمتى التباين والانحراف المعيارى؟

خطوات الحل :

ف	ك	س	س ك	س^{2}ك
٤-	٣	٦	١٨	١٠٨
٨-	٥	١٠	٥٠	٥٠٠
١٢-	٦	١٤	٨٤	١١٧٦
١٦-	٤	١٨	٧٢	١٢٩٦
٢٠-	١٠	٢٢	٢٢٠	٤٨٤٠
٢٤-	٧	٢٦	١٨٢	٤٧٣٢
٢٨ -٣٢	٥	٣٠	١٥٠	٤٥٠٠
مجـ	٤٠		٧٧٦	١٧١٥٢

$$\bar{س} = \frac{٧٧٦}{٤٠} = ١٩,٤ \text{ تقريباً}$$

$$ع^{٢} = \frac{١٧١٥٢}{٤٠} - (١٩,٤)^{٢}$$

$$= ٤٢٨,٨ - ٣٧٦,٣٦ = ٥٢,٤٤$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{١٧١٥٢}{٤٠} - (١٩,٤)^{٢}}$$

$$= \sqrt{٤٢٨,٨ - ٣٧٦,٣٦}$$

$$= \sqrt{٥٢,٤٤} = ٧,٢٤$$

وجدير بالذكر أنه لتبسيط العمليات الحسابية يمكن للباحث طرح وسط فرضى من بين مراكز الفئات.

مزايا الانحراف المعيارى

١ - يعتبر من أدق مقاييس التشتت

٢ - يعتبر أداة تحليلية قوية فى وصف خصائص التوزيع للمجتمع الأصلى.

٣ - يساعد الباحث على التنبؤ بما تكون عليه الحالات من توزيعات وتشتت مستقبلى فى القيم من خلال وحدات الانحراف المعيارى وما تعكسه من مؤشرات.

ولما كان الانحراف يتوقف على الوحدات المستخدمة فى قياس المشاهدات محل الدراسة لهذا فان قيمته المطلقة غير مناسبة لاغراض المقارنة. ومن ثم يمكن الاعتماد على معامل الاختلاف كمقياس للتشتت محرر من أثر الوحدات المستخدمة فى القياس ويمكن حساب معامل الاختلاف باستخدام المعادلة التالية.

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\text{الوسط الحسابى}} \times ١٠٠$$

مقاييس التشتت للمتغير المتقطع:

نسبة التباين The Variation Ratio

يماثل هذا المقياس فى بساطة وسهولة حسابة مقياس المدى. ويستخدم مقياس نسبة التباين فى حالة البيانات المبوبة The Grouped data ويلائم حالات المقاييس (التصنيفية) Nominal Scales .

تقيس نسبة التباين درجة تركز الحالات حول الفئة المنوالية The Modal Category بدلا من قياس التوزيع للقيم فى جميع الفئات.

وتحسب نسبة التباين من المعادلة الآتية :

$$\text{نسبة التباين}\,(ت_{ن}) = ١ - \frac{\text{عدد الحالات فى الفئة المنوالية}}{\text{المجموع الكلى لعدد الحالات}}$$

مـثال (٧) :

باستـخدام نسبـة التبـاين ، وضـح مدى تركز أو تشتت الحـالات حـول الفئـة المنواليـة فى توزيـع مـا لأحـد المتغيـرات. إذا كـان عـدد الحـالات حـول فى الفئـة المنوالية ثلاث حالات ، و المجموع الكلى للحالات عشرة حالات.

الحل :

$$\text{نسبة التباين (ت}_\text{ن}\text{)} = ١ - \frac{٣}{١٠} = ٠,٧$$

دليل التباين الكيفى :

يستخدم دليل التباين الكيفى Index of Qualitative variation للمقارنة بين التباين المشاهد للمتغير التصنيفى، والتباين الاقصى المتوقع . يتم حساب التباين المشاهد بحساب الاختلافات فى التوزيع . بينما يمثل التباين المتوقع أقصى تشتت يمكن أن يحدث لتوزيع معين. ثم يتم حساب دليل التشتت الكيفى (ف ت ك) من المعادلة الآتية:

$$\text{دليل التباين الكيفى} = \frac{\text{التباين المشاهد}}{\text{التباين المتوقع}} \times ١٠٠$$

مثال (٨) :

يشـتـمل الجـدول الآتى على عـدد المشـتـركين فى ندوتين علمـيـتين من تخصصـات علميـة هى الخدمه الاجتماعية، علم الاجـتمـاع، علم النفس، وعلم الانثرويولوجـيا. اضـافـة إلى التـباينـات المشاهدة والمتوقـعـة فى مـشاركـة هؤلاء الباحثين فى كل ندوة على حدة . والمطلوب قياس التباين باستخدام دليل التباين الكيفى.

التخصيص للمشاركين	الندوةالعلمية الأولى		الندوة العلمية الثانى	
	المشاهد	متوقع	المشاهد	متوقع
خدمة اجتماعية	٢	٥	٨	٧
علم اجتماع	١٧	٥	٦	٧
علم النفس	١	٥	٩	٧
الانثروبولوجيا	صفر	٥	٥	٧
	ن = ٢٠		ن = ٢٨	

الحل : نلاحظ فى هذا المثال أن المتغير هنا متقطع وليس متصل .

لحساب دليل التباين الكيفى من المعادلة

$$\text{دليل التباين} = \frac{\text{التباين المشاهد}}{\text{التباين المتوقع}} \times ١٠٠$$

١- كيف يتم حساب التباين المشاهد من الجدول:

نقوم بجمع حاصل ضرب المشاهدات لكل الازواج الممكنةمنها فى كل ندوة علمية على حدة.

أ - بالنسبة للندوه الأولى:

التبـاين المشـاهد = (٢ × ١٧) + (٢ × ١) + (٢ × صـفـر) + (١٧ × ١) + (١٧ × صفر) + (١ × صفر) = ٣٤ +٢ +١٧ = ٥٣

٢ - يتم حساب التباين الاقصى بجمع كل بيانات التوزيع المشاهد (٢٠) ثم قسمتها على عدد التصنيفات التخصصيه للمشاركين فى الندوة (٤) ثم يوزع الناتج من القسمة بالتساوى على كل مصنف. فالناتج من القسمة = $\frac{٢٠}{٤}$ = ٥ يصبح هو الرقم الدال على التباين المتوقع أمام كل قيمة مناظرة من التباين المشاهد.

٣ – حساب التباين المتوقع مثلما تم حساب التباين المشاهد. بأنه يساوى حاصل جمع كل زوجين احتمالين من التباين المتوقع.

التباين المتوقع = [(٥ × ٥) + (٥ × ٥) + (٥ × ٥)]

+ [(٥ × ٥) + (٥ × ٥)] + ٥ × ٥ = ١٥٠

∴ نسبة التباين المشاهد إلى المتوقع يمثل دليل التباين الكيفى

$$= \frac{٥٣}{١٥٠} \times ١٠٠ = ٣٥,٣٣ \ (١٠٠)$$

= ٣٥,٣٣ ٪

بالمثل يمكن تكرار الخطوات السابقة فى حساب دليل التباين الكيفى فى حالة الندوة العلمية الثانية وسوف نحصل على قيمة هذا الدليل وتساوى (٩٨,٣٠ ٪) نخلص من قيم دليل التباين الكيفى إلى أن النتائج فى هذا البحث مرضية نظراً لأن الندوة الأولى كان التشتت محدودا جدا فى توزيعها ومن ثم اعطت قيمة مئوية منخفضة للدليل بينما كان التشتت عاليا فى الندوة الثانية، لذلك كانت القيمة المئوية للدليل عالية.

من فوائد دليل التباين أنه مقياس مفيد للتباين فى توزيعات البيانات المتقطعة لأنه يقيم التشتت المشاهد داخل أى توزيع فى مقابل التوزيع المتوقع.

المفاهيم الأساسية Key Concepts

١ – التباين (التشتت) Dipersion : يعرف بكمية أو مقدار الإختلاف أو عدم التجانس فى أى توزيع للبيانات

٢ – دليل التباين الكيفى Index of qualitative Variation (IQV) يعرف بمقياس التشتت للمتغيرات المتقطعة التى يتم تنظيمها فى توزيعات تكرارية

٣ – الانحراف الربيعى يعرف بنصف المسافة بين الربيع الثالث والربيع الأول

٤ – الانحرف المعيارى للعينة Sample Standard Deviation (ع) يعرف بقيمة الجذر التربيعى للتباين.

٥ – تباين العينة (ع) Sample Variance يعرف بحاصل جمع كل انحرافات القيم عن الوسط الحسابى، وتربيعاتها.

تمـــــارين

١ - اكمل ما يأتى بالعبارات المناسبة والصحيحة

(أ) أن نسبة مقدار التباين الذى يتم ملاحظته فعليا فى أى توزيع للقيم إلى مـقـدار التـبـاين الأقـصى الممكن وجـوده فى هذا التـوزيع يعرف......................

(ب) يعـرف المدى بأنه .... بين أعلى وأقل القـيم فى التوزيع.

(جـ) يتـجنب الانحـراف الربيـعى مـشكلات القيم الشاذة باعـتـماده على............. فقط

(د) إن معادلة الانحراف المتوسط هى:......................

(هـ) يرتبط الانحراف المعيارى بالتباين حيث يمثل الأول .... للثانى.

(و) كلما كان التوزيع أكثر تشتتا ، فإن قيمة التباين المعيارى تكون ....

(ز) لو كان التوزيع غير ......بالتشتت ، يكون الانحراف المعيارى له..............

٢ - فيما يلى عدد من الاختيارات تحت كل عبارة والمطلوب وضع علامة (√) أمام الاختيار المناسب لهذه العبارة.

(أ) يستخدم الانحراف المعيارى تباين أو تشتت القيم من

(١) الوسط الحسابى

(٢) الوسيط

(٣) المنوال

(٤) التباين

(٥) جميع الاختبارات السابقة

(ب) ما المجموعة التى تعتبر اكثر تشتتا فى قيمها من المجَموعات التالية:

(١)	١٠	١١	١٢	١٥	١٨	٨٤	٩٦
(٢)	١٠	٤٢	٥٢	٤١	٤٦	٤٧	٩٦
(٣)	١٠	٢٢	٣٩	٤٠	٥٦	٨٥	٩٦
(٤)	١٠	١٠	١٠	٢٠	٢٠	٢٠	٩٦

(جـ) ما المقياس الذى لايرتبط بمقياس آخر من مقاييس التشتت الآتية:

(١) الانحراف المعيارى

(٢) المدى

(٣) التباين

(٤) المتوسط الحسابى

(د) أى مقياس من مقاييس التشتت الآتية لايعتمد على المقدار الصحيح لكل قيمة:

(١) الانحراف المعيارى

(٢) الانحراف الربيعى

(٣) التباين

(٤) لامقياس من المقاييس الثلاثة السابقة

(٥) كل المقاييس الثلاثة السابقة.

(هـ) إن المشكلة الرئيسية فى استخدام الانحراف الربيعى تتمثل فى :

(١) عدم الثبات

(٢) حمل فى حساباته

(٣) لا يتيح استخدام مستويات رياضية دقيقة فى الحسابات

(و) لو قورن المدى بالانحراف الربيعى يكون:

١ - أكبر

٢ - أصغر

٣ - سالب

٤ - كل الاختيارات الثلاثة السابقة

(ز) ما الكلمة التى تستخدم فى تعريف انتشار القيم حول مقياس للنزعة المركزية:

١ - المدى

٢ - التشتت

٣ - التوزيع

٤ - التباين

٥ - الاختيارات الاربعة السابقة.

٣ - فى دراسة إجتماعية أجريت على الحراك الوظيفى (الترقى) للرجال والنساء داخل أحد الأقسام فى تنظيم حكومى . وتوضح بيانات الجدول الآتى أن النساء تنتظر سنوات أطول فى درجتها الوظيفية عن نظائرهن من الرجال حتى يحصلن على ترقية لدرجة أعلى وأن التفرقة بسبب تباين النوع (الجندرة) : المطلوب حساب المتوسط والانحراف المعيارى للرجال والنساء . ثم أذكر رأيك حول مشكلة التفرقة على أسس النوع فى الترقى الوظيفى (الجندرة) من خلال ما تحصل عليه من اجابات .

عدد السنوات التى يتم قضاؤها فى الاداء الوظيفى قبل الترقية الأولى	عدد العاملين	
	ذكور	أناث
١	١٣	٤
٢	٢٥	١٨
٣	٢٠	١٤
٤	١٢	١٠
٥	٣	١٠

٤ – فى دراسة أجريت على أحد السجون فى مصر للتعرف على عدد الذين يتم العفو عنهم لحسن السير والسلوك ولأسباب أخرى قبل قضائهم مدة العقوبة القانونية. والعلاقة بين عدد هؤلاء وعدد مرات القاء القبض على المتهمين من جانب الشرطة . أعطت الدراسة البيانات الموضحة بالجدول الآتى والمطلوب حساب المتوسط لهذه العينة.

عدد مرات القاء القبض على المتهمين	عدد الذين تم العفو عنهم من المسجونين
صفر	١٨
١	٢٢
٢	١٤
٣	٢٠
٤	١٣
٥	٧
٦	٦
	١٠٠

٥ – أجريت دراسة مسحية للتعرف على عدد الناخبين المسجلين أسمائهم فى أحدى الدوائر الانتخابية خلال خمس دورات انتخابية لمجلس الشعب المصرى. وأعطت نتائج المسح النسب المئوية لمن أدلوا بأصواتهم فى

هذه الانتخابات من الذكور والأناث كما هو مبين أدناه والمطلوب معرفة وحساب:

(أ) هل كان التصويت الفعلى للإناث اللاتى سجلن اصواتهن أكثر من التصويت الفعلى للرجال فى الدورات الانتخابية الخمس؟

(ب) احسب الانحراف المعيارى لكل من الذكور والأناث

النسبة المئوية للأصوات المسجلة
لمن شـارك من الناخبين فى انتخـابات
مجـلس الشـعب

النوع	١٩٧٥	١٩٨٠	١٩٨٥	١٩٩٠	١٩٩٥
	الدورات الأنتخابية (بالسنوات) لمجلس الشعب المصرى				
الأنـــاث					
أرامل	٥٨	٥٦	٦٠	٥٨	٥٩
متزوجات وغير عاملات	٤٩	٤٠	٤٩	٤١	٥٨
متزوجات وعاملات	٥٥	٥٧	٥٩	٥٣	٥٨
الذكـــور					
عزاب	٦٧	٦٠	٦٨	٥٦	٦٥
متزوجون ولا يعملون	٤٠	٤٠	٤٢	٤٥	٤٣
متزوجون ويعملون	٦٠	٥٨	٦٢	٥٥	٦٠

النسبة المئوية الاجمالية للناخبين المسجلين

إنـــاث		ذكـــور	
ارامل	٣٥	عزاب	٣٦
متزوجات وغير عاملات	٣٣	متزوجون ولا يعملون	١٠
متزوجات وعاملات	٣٢	متزوجون ويعملون	٥٤
	١٠٠		١٠٠

٦ - احسب قيمة التباين للقيم الآتية :

١٠ ، ١٥ ، ١٨ ، ٢١ ، ٢٤ ، ٢٩ ، ٣٢ ، ٣٣

الفصل السادس
الارتباط والانحدار الخطى

أنـــواع الارتباط :

١ - الارتباط الخطى البسيط :

(أ) معامل بيرسون .

(ب) معامل سيبرمان .

(جـ) معامل فاى .

(د) معامل التوافق .

٢- الارتباط الجزئى والمتعدد .

الانحدار الخطى .

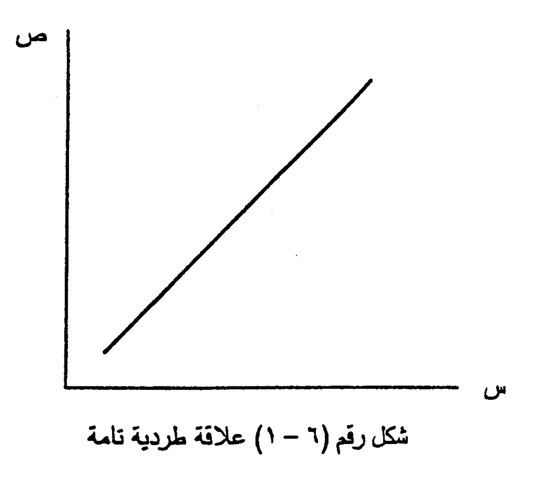
الفصل السادس

الارتبـــاط والانحــــدار

مقــدمة :

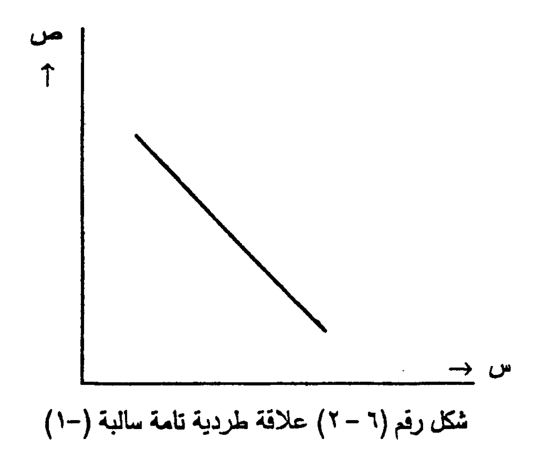
إذا كنا فى الفصول السابقة قد عرضنا لعدد من البيانات الإحصائية المتعلقة بتوزيع ووصف متغير واحد سواء من خلال الأشكال البيانية أو قياس هذا المتغير بواسطة عدد من المقاييس الإحصائية مثل المتوسط الحسابى، الوسيط، المنوال وأيضاً مقاييس التشتت، فإن الارتباط كأحد مجالات الإحصاء التطبيقية يهم جميع المهتمين بدراسة العلاقة بين الظواهر المختلفة فى مختلف فروع المعرفة من علوم رياضية، تجارية، تربوية، نفسية، اجتماعية، وغيرها من فروع العلوم الأخرى. ففى العلوم الرياضية نجد من خلال النظريات الرياضية عدداً من القوانين تحكم العلاقة بين المتغيرات بشكل واضح ، مثال ذلك القوانين التى تحدد العلاقة بين طول الضلع ومساحة الشكل الهندسى المنتظم مثل المربع، المستطيل والمثلث، كذلك العلاقة بين المساحة وطول نصف القطر للدائرة وهكذا. وهذا النوع الواضح من العلاقة بين متغيرين نطلق عليه ارتباطاً تاماً موجباً بمعنى أنه كلما زاد طول نصف قطر الدائرة لابد أن تزيد معه المساحة والعكس صحيح تماماً. كذلك نجد فى العلوم الطبيعية أمثلة للعلاقة التامة الكاملة بين متغيرين كالعلاقة بين حجم الغاز وضغطه حيث يحكم تلك العلاقه قانون محدد. فبغرض ثبوت درجة الحرارة،نجد علاقة عكسية بين ضغط الغاز وحجمه أى كلما زاد الضغط قل الحجم والعكس صحيح أيضاً.

وإذا أردنا أن نعبر رياضياً عن نوع العلاقة الارتباطية التامة فى المثالين السابقين، نجد أن العلاقة الأولى التامة موجبة الإشارة (+١) بينما نجدها تامة سالبة الإشارة فى المثال الثانى عند القيمة (-١) . هذا ويمكن بالتمثيل البيانى أن نفرق بين هذين النوعين من الارتباط التام فالنوع الأول وهو الارتباط الموجب نجد العلاقة بين المتغيرين خطية تصاعدية تبدأ من نقطة الأصل وتتجه بقيم متزايدة ناحية اليمين. كما يتضح من الشكل رقم (٦-١) .



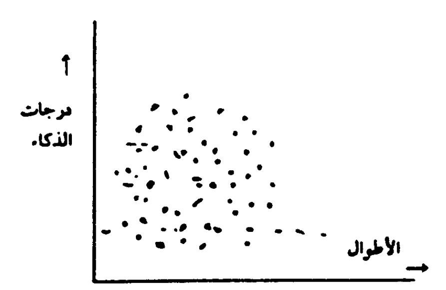
شكل رقم (٦ – ١) علاقة طردية تامة

كذلك يمكن رسم العلاقة الارتباطية التامة السالبة وقيمتها (–١) إذا ما اتبعنا نفس الخطوات السابقة فى الرسم البيانى فنحصل على شكل انتشارى تمثل معظم نقاطه خطاً مستقيماً يبدأ بقيم أصغر للمتغيرين (س،ص) من ناحية اليمين ثم يتجه تصاعدياً بقيم متزايدة لقيم (ص) وقيماً تناقصية مناظرة للمتغير (س) كما يتضح من الشكل رقم (٦–٢).



شكل رقم (٦ – ٢) علاقة طردية تامة سالبة (–١)

أما فى العلوم التربوية والاجتماعية والنفسية والتى تتعامل مع الإنسان كفرد يتصف بتباين فى الميول والنزعات والتى تنعكس على تصرفاته بقدر هذا التنوع. فقد يتعذر تماماً تحقيق هذا النوع من الارتباط التام. ولا تقتصر كل حالات المتغيرات على وجود علاقة ارتباطية سالبة أو موجبة الاتجاه فيما بينها، فقد تنعدم العلاقة. فمثلاً لم يتوصل العلماء إلى وجود علاقة بين طول الفرد ودرجة ذكائه، ونصف هذه الحالة من عدم الارتباط مجازاً بالارتباط الصفرى. وإذا حاولنا تمثيل الارتباط الصفرى بيانياً باستخدام الشكل الانتشارى فسوف نجد أن النقط شديدة التشتت والتنافر تتنافى معها تماماً وجود أشكال خطية أو انحنائية كما يوضح ذلك الشكل رقم (٦-٣).



شكل رقم (٦ - ٣) ارتباط صفرى

أما بالنسبة لأشكال الارتباط بين متغيرين فهى أما أن تكون خطية أو انحنائية تبعاً لدرجة العلاقات بينهما. فالعلاقة الخطية تتبع معادلة من مجموعة معادلات الدرجة الأولى على النحو التالى: ص = أ +ب س. أى يكون المتغير(ص) دالة للمتغير (س) أما العلاقة الانحنائية فتمثلها معادلات من الدرجة الثانية والثالثة.

الارتبـــــاط

ينقسم الارتباط إلى الأنواع التالية:

١ – الارتباط الخطى البسيط Simple Linear Correlation ويقيس معامل الارتباط العلاقة بين متغيرين أحدهما معتمد (ص) والآخر مستقلاً (س).

٢ – الارتباط الجزئى Partial Correlation ويشير إلى العلاقة الارتباطية بين ظاهرتين مع إبقاء العوامل الأخرى ثابتة.

٣ – الارتباط المتعدد Multiple Correlation يستخدم لقياس العلاقة بين متغير تابع ومتغيرات أخرى مستقلة فى وقت واحد.

١-الارتباط البسيط ومعاملاته هى :

أ – معامل بيرسون.

ب– معامل سبيرمان.

جـ– معامل فاى.

د– معامل التوافق.

حساب معامل بيرسون للارتباط « ر » من القيم الخام:

تقوم طريقة بيرسون فى حساب معامل الارتباط بين متغيرين (س ، ص) على استخدام انحرافات القيم عن المتوسط الحسابى لكل منهما، أى الفرق بين (س – سَ) للمتغير الأول، (ص – صَ) للمتغير الثانى. وذلك على أساس أن الارتباط يقيس العلاقة بين التغير فى قيم (س) والتغير فى قيم (ص). وتعتبر طريقة قياس انحرافات القيم عن متوسطها الحسابى هى أفضل طرق لقياس هذا التغير وتحسب قيمة معامل الارتباط بين المتغيرين (س)، (ص) من العلاقة الآتية بفرض أن ن= حجم العينة.

$$ر = \frac{\frac{1}{ن} مجـ س ص - \bar{س}\bar{ص}}{\sqrt{\left[\frac{1}{ن} مجـ س^2 - \bar{س}^2\right]} \cdot \sqrt{\left[\frac{1}{ن} مجـ ص^2 - \bar{ص}^2\right]}}$$

معادلة رقم (٦ - ١)

$$\because \frac{مجـ س}{ن} = \bar{س} \quad ، \quad \frac{مجـ ص}{ن} = \bar{ص}$$

∴ يمكن صياغة المعادلة (٦ - ١) لتصبح :

$$ر = \frac{مجـ س ص - \frac{(مجـ س)\ (مجـ ص)}{ن}}{\sqrt{\left[مجـ س^2 - \frac{(مجـ س)^2}{ن}\right]} \cdot \sqrt{\left[مجـ ص^2 - \frac{(مجـ ص)^2}{ن}\right]}}$$

معادلة رقم (٦ - ٢)

وللتخلص من المقادير الكسرية فى المعادلة رقم (٦-٢) بضرب البسط والمقام × ن نحصل على المعادلة الآتية:

$$ر = \frac{ن مجـ س ص - (مجـ س)\ (مجـ ص)}{\sqrt{\left[ن مجـ س^2 - (مجـ س)^2\right] \cdot \left[ن مجـ ص^2 - (مجـ ص)^2\right]}}$$

معادلة رقم (٦ - ٣)

مثال رقم (١) :

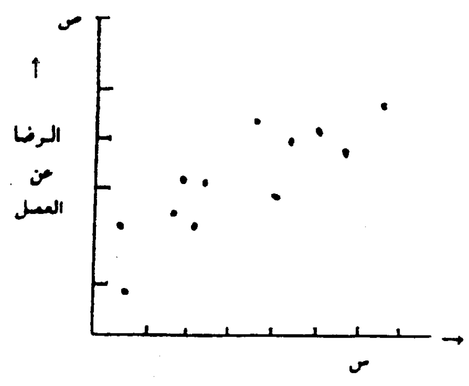
يوضح الجدول الآتى توزيع الدخل اليومى لعينة مكونة من اثنى عشر عاملاً وأيضاً درجاتهم فى الرضا عن العمل. والمطلوب معرفة حجم العلاقة واتجاهها؟

العمال	الدخل اليومى (س)	الرضا عن العمل (ص)
١ - محمد	١٠,٥٠	٩٤
٢ - أحمد	٩,٥٠	٨٩
٣ - علية	٩,٠٠	٨٩
٤ - حسين	٨,٢٥	٩٠
٥ - منال	٨,٠٠	٨٤
٦ - زينب	٧,٥٠	٩٢
٧ - ماهر	٦,٢٥	٨٦
٨ - على	٦,٠٠	٨١
٩ - ولاء	٥,٧٥	٨٦
١٠- طارق	٥,٥٠	٨٢
١١- فاطمة	٤,٥٠	٧٤
١٢- حامد	٤,٢٥	٨١

الحـــل :

يمكن أولاً توضيح شكل العلاقة بين المتغيرين الدخل (س) والرضا عن العمل (ص) باستخدام الشكل الانتشارى لتوزيع الحالات الإثنى عشر وذلك بتوقيع نقطة تمثل كل حالة باحداثيها (س،ص) فمثلاً فى الحالة الأولى نجد أن قيمة دخل محمد (١٠,٥٠) وهى تمثل قيمة (س) يناظرها قيمة (ص) (٩٤) وهكذا نوقع النقط فنحصل على الشكل الانتشارى رقم (٦-٤) الذى يوضح نزعة معظم النقط حول خط مستقيم تصاعدى يبدأ من ناحية نقطة الأصل ويتجه يميناً وهذا يدل

على وجود علاقة ارتباطية خطية طردية بمعنى كلما زاد الأجر زاد الرضا عن العمل.



شكل رقم (٦ - ٤) توزيع انتشارى للأجور والرضا عن العمل

وللحصول على معامل الارتباط باستخدام المعادلة رقم (٦-٣).

فإننا نقوم بحساب قيم (مجـ س ص) ، (مجـ ص) ، (مجـ س٢) ، (مجـ ص٢) كما هو موضح فى جدول رقم (٦-١).

جدول رقم (٦-١)

العمال	الدخل س	الرضاعن العمل ص	س٢	ص٢	س ص
١	١٠,٥٠	٩٤	١١٠,٢٥	٨٨٣٦	٩٨٧,٠٠
٢	٩,٥٠	٨٩	٩٠,٢٥	٧٩٢١	٨٤٥,٥٠
٣	٩,٠٠	٩١	٨١,٠٠	٨٢٨١	٨١٩,٠٠
٤	٨,٢٥	٩٠	٦٨,٠٦	٨١٠٠	٧٤٢,٥٠
٥	٨,٠٠	٨٤	٦٤,٠٠	٧٠٥٦	٦٧٢,٠٠
٦	٧,٥٠	٩٢	٥٦,٢٥	٨٤٦٤	٦٩٠,٠٠
٧	٦,٢٥	٩٦	٣٩,٠٦	٧٣٩٦	٥٣٧,٥٠

العمـــال	الدخل س	الرضاعن العمل ص	س٢	ص٢	س ص
٨	٦,٠٠	٨١	٣٦,٠٠	٦٥٦١	٤٨٦,٠٠
٩	٥,٧٥	٨٦	٣٣,٠٦	٧٣٩٦	٤٩٤,٥٠
١٠	٥,٥٠	٨٢	٣٠,٢٥	٦٧٢٤	٤٥١,٠٠
١١	٤,٥٠	٧٤	٣٠,٢٥	٥٤٧٦	٣٣٣,٠٠
١٢	٤,٢٥	٨١	١٨,٠٦	٦٥٦١	٣٤٤,٢٥
مجـ	٨٥	١٠٣٠	٦٤٦,٤٩	٨٨٧٧٢	٧٤٠٢,٢٥

وبالتعويض فى المعادلة رقم (٦-٣) لإيجاد قيمة ر: فإن:

$$ر = \frac{ن\ مجـ\ س\ ص\ -\ (مجـ\ س)\ (مجـ\ ص)}{\sqrt{\left[ن\ مجـ\ س^{٢} - (مجـ\ س)^{٢}\right] . \left[ن\ مجـ\ ص^{٢} - (مجـ\ ص)^{٢}\right]}}$$

$$ر = \frac{١٢\ (٧٤٠٢,٢٥) - (٨٥)\ (١٠٣٠)}{\sqrt{\left[١٢\ (٦٤٦,٤٩) - (٨٥)^{٢}\right] . \left[١٢\ (٨٨,٧٧٢) - (١٠٣٠)^{٢}\right]}}$$

$$= ٠,٨٣٧$$

وإذا كانت قيمة (ر) تدل على وجود علاقة طردية قوية بين الأجر اليومى والرضا عن العمل، فإن (ر) لا تصف تلك العلاقة بمعنى ، هل العلاقة الطردية بين المتغيرين ترجع إلى عامل الصدفة البحتة إم إنها علاقة جوهرية ، ولمعرفة ذلك يجب قياس الدلالة الإحصائية لتلك العلاقة حتى يمكن الحكم عليها.

حساب الدلالة الإحصائية لمعامل الارتباط:

ولحساب الدلالة الإحصائية لمعامل الارتباط فى المثال السابق يجب اولا توضيح ماذا تعنى الدلالة الإحصائية Statistical Significance ثانيا توضيح

الدلالة الإحصائية من خلال خطوات اختبار الفرض الصفرى* ويرمز له بالرمز (H_0) .

تشير الدلالة الإحصائية إلى الحد الاقصى لاحتمال وقوع الباحث فى الخطأ من النوع الاول، ويطلق على هذا الخطأ خطأ الفا ويرمز له بالرمز اللاتينى (α) (ويستخدم نفس الرمز للاشارة إلى مستوى الدلالة) . ويحدث هذا النوع من الخطأ عندما يرفض الفوض الصغرى (H_0) فى حين أنه صحيح. وتعتبر نسبة (٥ ٪) حد أقصى مقبولا للوقوع فى مثل هذا الخطأ . ويعنى مستوى الدلالة (٠,٠٥) أنه إذا كررنا التجربة (١٠٠) مرة فمن المحتمل أن نرفض الفرض الصفرى (٥) مرات وهو فى الواقع صحيح .فدرجة الشك فى النتيجة (٥ ٪) ودرجة الثقة تصل إلى (٩٥ ٪) وعلى الباحث تحديد مستوى الدلالة قبل جمع بياناتة من الميدان. وقد يختار (٠,٠٥) ، (٠,٠١) ، (٠,٠١)

- حساب مستوى الدلالة من خلال اختيار الفرض الصفرى (H_0) .

خطوات الحل :

١- صياغة الفرض الصفرى H_0 .

يمكن صياغة الفرض الصفرى من خلال الصياغات الثلاثة التالية :

(أ) H_0 العلاقة بين س ، ص = صفر

(ب) H_0 العلاقة بين الدخل والرضا عن العمل = صفراً

صياغة أخرى

(جـ) H_0 لا توجد علاقة دالة احصائيا بين الدخل والرضا عن العمل.

ولكل فـرض صـفرى فـروض بديلة وقـد يكـون هذا الفـرض البـديل إتجاهيا Directional او غير اتجاهى Nondirectional

(*) انظر فصل اختبار الفروض في اعتماد علام ويسري رسلان، أساسيات الإحصاء الاجتماعي، دار قطري بن الفجاءة للنشر والتوزيع، الدوحة ١٩٩١، ص ص ٢٢١-٢٦٨، زكريا الشربيني، الإحصاء اللابارامتري في العلوم النفسية والتربوية والاجتماعية، مكتبة الانجلو المصرية، القاهرة، ١٩٩٠ ص ص ٥٧-٦٦.

٢ - صياغة الفرض البديل الاتجاهى

H_1 العلاقة بين الدخل والرضا عن العمل > الصفر

H_2 العلاقة بين الدخل والرضا عن العمل < الصفر

٣ - صياغة الفرض البديل غير الاتجاهى

H_1 العلاقة بين الدخل والرضا عن العمل ≠ صفر

٤ - حساب قيمة (ر) من المعطيات الامبيريقية

٥ - تحديد درجات الحرية ورمزها (د.ج)

د.ج = ن - ٢ حيث ن تشير لحجم العينة

٦ - استخدام جدول توزيع القيم الحرجة لمعامل الارتباط:

حيث أن قيمة ر المحسوبة = ٨٣٧٤ ر

ومن ملحق توزيعات القيم الحرجة رقم (١) عند مستوى دلالة (٠,٠١) ودرجات حرية = ١٠

∴ ر= ٠,٨٠٨

ر المحسوبة = ٠,٨٣٧٤ › ر المجدولة = ٠,٨٠٨

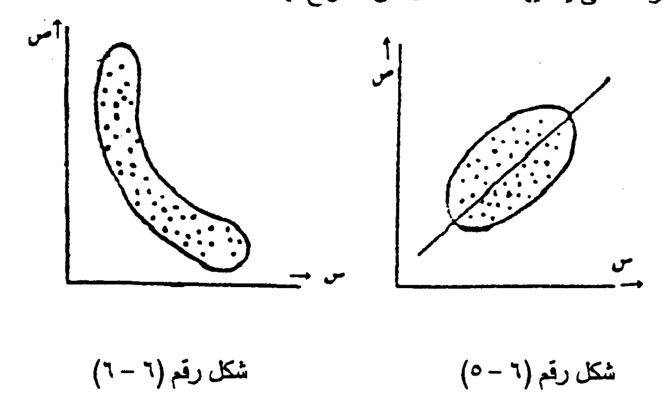
نخلص من ذلك إلى رفض الفرض الصفرى ومن ثم توجد علاقة جوهرية بين المتغيرين عند مستوى الدلالة (٠,٠١).

العوامل المؤثرة على حجم معامل الارتباط (ر) :

١- العلاقة الخطية بين المتغيرين (س،ص):

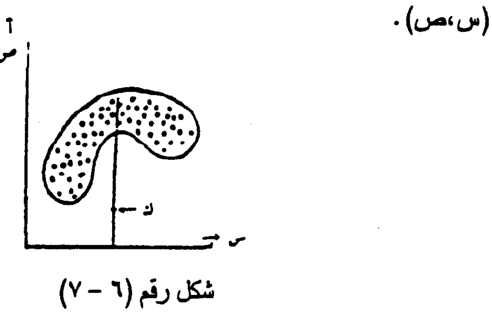
كما ذكرنا سابقاً - تعتبر العلاقة بين المتغيرين (س،ص) أحد أنواع العلاقات بل وأبسطها سواء كانت تلك العلاقة الخطية سالبة أو موجبة كما تعتبر العلاقة الخطية بين المتغيرين (س،ص) كما يتضح ذلك من الشكل رقم (٦-٥) أفضل أنواع العلاقات لتحقيق الدقة فى حساب قيمة معامل الارتباط. وكلما كانت العلاقة بين المتغيرين أكثر اتجاهاً للشكل المنحنى كما يتضح ذلك من الشكلين رقم (٦-٦) ورقم (٦-٧) نجد أن معامل الارتباط لا يعطى المؤشرات المتوقعة للعلاقة

بل كلما زادت درجة الإنحناء فى الشكل الانتشارى للعلاقة بين المتغيرين قد تقل المؤشرات التى يعطيها هذا المعامل عن المتوقع لها.



شكل رقم (٦ - ٦)　　　　شكل رقم (٦ - ٥)

ففى الشكل رقم (٦-٧) نجد التوزيع انحنائياً للعلاقة بين المتغيرين (س،ص). حيث يكون هناك زيادة تأخذ الشكل الإنحنائى حتى النقطة (ك) ثم تبدأ العلاقة فى الانحدار بين المتغيرين. ونجد لهذا الشكل أمثلة متعددة فى حياتنا اليومية منها،على سبيل المثال ، العلاقة بين التوتر والأداء فقد تتزايد العلاقة بينهما طردياً إلى حد ما مثلاً عند النقطة (ك) ثم تبدأ العلاقة فى أخذ شكل آخر يخالف الأول وهى التى تمثلها المساحة بعد النقطة (ك) . فحساب معامل الارتباط بين المتغيرين سوف تنقصه الدقة للتفاوت الواضح فى التشتت للقيم وتباين الاتجاهات (سالب وموجب) للقيم وأحداثياتها (س،ص). بالمثل نجد افتقاد الدقة فى حساب معامل (ر) فى الشكل الانتشارى رقم (٦-٦) للعلاقة الإنحنائية بين المتغيرين (س،ص).



شكل رقم (٦ - ٧)

من ثم يصلح معامل بيرسون لتفسير العلاقات الخطية بين متغيرين حيث أن استخدام العلاقات الإنحنائية تقلل من دقة معامل الارتباط كما تجعل المؤشرات التى يعطيها أقل مما هو متوقع منها.

٢- تجانس المجموعة :

يعتبر تجانس المجموعة مؤثراً هاماً على حجم معامل الارتباط (ر) فكلما ازداد تجانس متغير منهما أو كليهما يقل التباين.

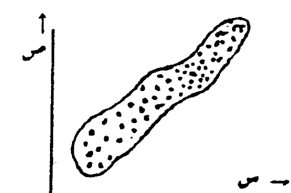
كلما كانت المجموعة أكثر تجانساً للمتغير الواحد أو لكليهما تقل القيمة المطلقة لمعامل الارتباط.

٣ - حجم المجموعة :

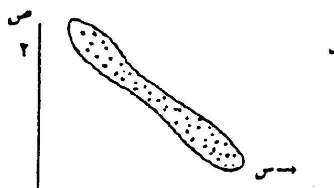
تتأثر دقة معامل الارتباط (ر) وفق كبر أو صغر حجم العينة ومن جهة أخرى لا تتأثر قيمة المعامل بهذا الحجم.

استخدام الشكل الانتشارى للعلاقة بين متغيرين (س، ص) فى فهم معنى معامل الإرتباط بينهما:

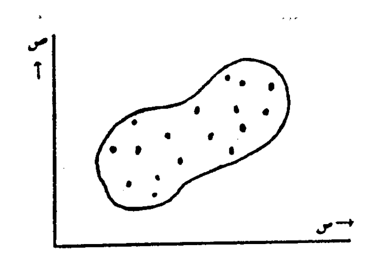
بعد حساب قيمة معامل الإرتباط من المعادلة رقم (٦ – ٣) فيمكن باستخدام الشكل الانتشارى فى تمثيل هذه المعادلة بيانياً فهم معنى قيمة معامل الارتباط، حيث يمثل دالة لميل slope المسار الذى تأخذه معظم القيم فى الشكل الانتشارى. وكلما تزداد مساحة شكل المنحنى تقل درجة العلاقة، بينما تحدد اتجاهها. ويتضح ذلك من الأشكال التالية أرقام (٦–٨ ، ٩ ، ١٠، ١١).



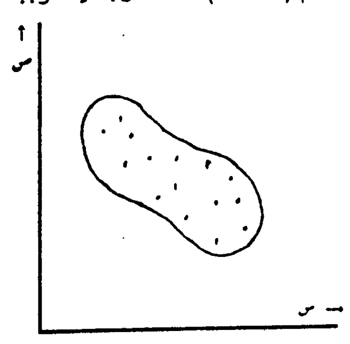
شكل رقم (٦ – ٨) علاقة ارتباطية موجبة عالية



شكل رقم (٦ – ٩) علاقة ارتباطية سالبة عالية



شكل رقم (٦ – ١٠) علاقة ارتباطية موجبة ضعيفة



شكل رقم (٦ – ١١) علاقة ارتباطية سالبة ضعيفة

جدول العلاقة بين حجم معامل الارتباط (حدود تقريبية) ودرجة العلاقة الارتباطية بين متغيرين (س،ص):

حجم ر	درجة العلاقة الارتباطية
(من ٠,٧٥ إلى ١,٠٠) ، (من -٠,٧٥ الى -١,٠٠)	علاقة ارتباطية عالية جداً إيجابية وسلبية
(من ٠,٥٠ إلى ٠,٧٤) ، (من -٠,٥٠ إلى - ٠,٧٤)	علاقة ارتباطية عالية إيجابية وسلبية
(من ٠,٢٥ إلى ٠,٤٩) ، (من -٠,٢٥ إلى -٠,٤٩)	علاقة ارتباطية متوسطة إيجابية وسلبية
(من صفر إلى ٠,٢٤) ، (من صفر إلى -٠,٢٤)	علاقة ارتباطية ضعيفة إيجابية وسلبية

هذا ويجدر التنويه إلى أن معامل الارتباط- يستخدم فى البيانات الترتيبيه للمتغيرات.

ويتسم معامل الارتباط بخاصيتين يستفاد منهما فى حسابه:

الأولى هى إذا طرحنا أو (جمعنا) رقم ثابت من جميع قيم (س) وثابت آخر من جميع قيم (ص) فإن قيمة معامل الارتباط لا تتغير. أما الخاصية الثانية فتشير إلى أن قيمة (ر) لا تتغير إذا قسمنا أو (ضربنا) جميع قيم (س) على ثابت وأيضاً جميع قيم (ص) على ثابت آخر.

ولتوضيح ذلك سوف نعطى مثالاً على الخاصية الأولى وهى طرح وسط فرضى من قيم (س) ثم وسط فرضى آخر من قيم (ص).

مـــثال (٢) :

يوضح الجدول الآتى قيم (س) و(ص) والمطلوب حساب معامل الارتباط من القيم الأصلية ثم استخدام (٣٩) كوسط فرضى لقيمم (س) ، (٣٥) كوسط فرضى لقيم الظاهرة (ص).

س	٢٥	٤٢	٣٥	٣٧	١٥	٢٤	٤٣	٥٣	٤٧	٣٩
ص	٢٢	٢٧	٤٥	٣٥	٣٣	٣٠	٣٢	٤٤	٤٥	٢٧

(أ) حساب معامل الارتباط (ر) من القيم الأصلية باستخدام المعادلة رقم (٦-٣):

قيم س	قيم ص	$س^2$	$ص^2$	س ص
٢٥	٢٢	٦٢٥	٤٨٤	٥٥٠
٤٢	٢٧	١٧٦٤	٧٢٩	١١٣٤
٣٥	٤٥	١٢٢٥	٢٠٢٥	١٥٧٥
٣٧	٣٥	١٣٦٩	١٢٢٥	١٢٩٥
١٥	٣٣	٢٢٥	١٠٨٩	٤٩٥
٢٤	٣٠	٥٧٦	٩٠٠	٧٢٠
٤٣	٣٢	١٨٤٩	١٠٢٤	١٣٧٦

	قيم س	قيم ص	$س^2$	$ص^2$	س ص
	٥٣	٤٤	٢٨٠٩	١٩٣٦	٢٣٣٢
	٤٧	٤٥	٢٢٠٩	٢٠٢٥	٢١١٥
	٣٩	٢٧	١٥٢١	٩٢٧	١٠٥٣
مجـ	٣٦٠	٣٤٠	١٤١٧٢	١٢١٦٦	١٢٦٤٥

$$ر = \frac{١٠(١٢٦٤٥) - (٣٦٠)(٣٤٠)}{\sqrt{\left[١٠(١٤١٧٢) - (٣٦٠)^٢\right] . \left[١٠(١٢١٦٦) - (٢٤٠)^٢\right]}}$$

ر = ٠,٤٧

(ب) استخدام الوسط الفرضى

س	ص	ح س س − ٣٩	ح ص ص − ٣٥	$ح^2_س$	$ح^2_ص$	$ح_س ح_ص$
٢٥	٢٢	− ١٤	− ١٣	١٩٦	١٩٦	١٨٢
٤٢	٣٧	٣	− ٨	٩	٦٤	− ٢٤
٣٥	٤٥	− ٤	١٠	١٦	١٠٠	− ٤٠
٣٧	٣٥	− ٢	صفر	٤	صفر	صفر
١٥	٣٣	− ٢٤	− ٢	٥٧٦	٤	٤٨
٢٤	٣٠	− ١٥	− ٥	٢٢٥	٢٥	٧٥
٤٣	٣٢	+ ٤	− ٣	١٦	٩	− ١٢
٥٣	٤٤	١٤	٩	١٩٦	٨١	١٢٦
٤٧	٤٥	٨	١٠	٦٤	١٠٠	٨٠
٣٩	٢٧	صفر	− ٨	صفر	٦٤	صفر
مجـ		− ٣٠	− ١٠	١٣٠٢	٦١٦	٤٣٥

$$\bar{\text{ح}}_{\text{س}} = \frac{\text{مجـ ح}_{\text{س}}}{\text{ن}} = -\frac{٣٠}{١٠} = -٣$$

$$\bar{\text{ح}}_{\text{ص}} = \frac{\text{مجـ ح}_{\text{ص}}}{\text{ن}} = \frac{-٣٠}{١٠} = -١$$

$$\text{ع}_{\text{س}} = \frac{\text{مجـ ح}^{٢}_{\text{س}}}{\text{ن}} - (\bar{\text{ح}}_{\text{س}})^{٢}$$

حيث $\text{ع}_{\text{س}}$ هى الإنحراف المعيارى (س)

$$= \sqrt{\frac{١٣٠٢}{١٠} - (-٣)^{٢}} = ١١٫٠١$$

$$\text{ع}_{\text{ص}} = \sqrt{\frac{\text{مجـ ح}_{\text{ص}}}{\text{ن}} - (\bar{\text{ح}}_{\text{ص}})^{٢}}$$

حيث $\text{ع}_{\text{ص}}$ = الانحراف المعيارى (ص)

$$= \sqrt{\frac{٦١٦}{١٠} - (-١)^{٢}}$$

$$= \sqrt{٦١٫٦ - ١} = \sqrt{٦٫٦٠} = ٧٫٧٨$$

وباستخدام المعادلة الآتية يحسب معامل الارتباط بطرح ثابت من قيم (س)، ثابت آخر من قيم (ص) (وسط فرضى) .

$$\text{ر} = \frac{\frac{\text{مجـ ح}_{\text{س}} \cdot \text{ح}_{\text{ص}}}{\text{ن}} - \bar{\text{ح}}_{\text{س}} \cdot \bar{\text{ح}}_{\text{ص}}}{\text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}$$

$$= \frac{\frac{٤٣٥}{١٠} - (-٣ \times -١)}{١١,٠١ \times ٧,٧٨}$$

$$= \frac{٤٣,٥٠ - ٣}{١١,٠١ - ٧,٧٨}$$

$$= \frac{٤٠,٥٠}{٨٥,٦٦} = ٠,٤٧$$

وهى نفس القيمة التى حصلنا عليها باستخدام القيم الأصلية:

استخدام طريقة الإنحرافات فى حساب معامل الارتباط للقيم المبوبة (طريقة بيرسون):

لا تختلف طريقة حساب معامل الارتباط فى حالة الجدول المزدوج لكل من المتغيرين (س،ص) عنها فى حالة حسابه للقيم غير المبوبة. فكلما كان عدد القيم كبيراً تزداد العمليات الحسابية لقيمة (ر) تعقيداً وصعوبة وإرهاقاً للباحث مالم يستخدم الحاسب الآلى - حتى فى هذه الحالة لا بد من إعداد مدخلات الحاسب من علاقات كثيرة تستنفذ قدراً من طاقة الباحث، وأن يقل كثيراً عن المبذول فى الطريقة اليدوية. ولتبسيط العمليات الحسابية فإننا ندخل التكرارت والاعتماد على قيم الإنحرافات بين مركز الفئة والمتوسط الحسابى أو بين المتوسطين الفرضى والواقعى للبيانات المبوبة للعلاقة بين المتغيرين (س،ص).

وتتلخص العمليات الحسابية لتقدير معامل بيرسون وفى الجدول المزدوج فى الخطوات التالية:

١ - من القيم المبوبة للمتغيرين (س،ص) يقوم الدارس بعمل جدولين هامشيين الأول يتعلق بقيم (س) والثانى بقيم المتغير (ص). ولعمل جدول هامشى يتم أخذ الصف الأفقى الأول بجميع بياناته مع الصف الأخير (صف المجموع) بجميع قيمة من الجدول الأصلى. فلو كانت الصفوف تمثل قيم المتغير (س) يصبح

هذا الجدول هامشى للمتغير (س) . وبالنسبة للمتغير (ص)، يتم إنشاء جدول هامشى أيضاً وذلك بأخذ العمود الأول (يميناً) مع العمود الأخير (عمود المجموع) بجميع قيمهما وبنفس توزيعهما فى الجدول الأصلى. ويسمى هذا الجدول الهامشى للمتغير (ص) بعد ذلك يقوم الدارس بعمل خمس خانات إضافية على كل جدول هامشى وذلك لحساب القيم المطلوبة لحساب (ر) من المعادلة رقم (٦-٤) ففى العمود الثانى يتم وضع التكرارت (ك) للمتغير (ص)، والعمود الثالث مركز الفئة بافتراض أنه يمثل المتوسط الفرضى والذى يناظره قيمة إنحراف تساوى صفراً. وفى العمود الثالث قيم الإنحراف (ص)، العمود الرابع قيم الإنحراف الثانى (ص). وفى العمود الخامس يتم حساب قيم حاصل ضرب ص ًك أما العمود الأخير وهو السابع من الجدول فيتم حساب قيم (ص٢ ك) . وبالمثل تكون جميع خانات الجدول الهامشى للمتغير (ص).

٢ - من معادلة (ر) رقم (٦ - ٤) :

$$\text{ر} = \frac{\left(\frac{\text{مجـ س' ص' ك}}{\text{مجـ ك}}\right) - \left(\frac{\text{مجـ س' ك}}{\text{مجـ ك}}\right) \times \left(\frac{\text{مجـ ص' ك}}{\text{مجـ ك}}\right)}{\sqrt{\left[\left(\frac{\text{مجـ س'}^{2}\text{ ك}}{\text{مجـ ك}}\right) - \left(\frac{\text{مجـ س' ك}}{\text{مجـ ك}}\right)^{2}\right] \times \left[\left(\frac{\text{مجـ ص'}^{2}\text{ ك}}{\text{مجـ ك}}\right) - \left(\frac{\text{مجـ ص' ك}}{\text{مجـ ك}}\right)^{2}\right]}}$$

يتضح إمكانية حساب جميع المقادير فى العلاقة السابقة باستثناء المقدار الأول من البسط وهو (مجـ س ً ص ً ك).

٣- لحساب قيمة المقدار (مجـ س ً ص ً ك) نقوم بعمل جدول ثالث يتضمن التكرارت الأصلية للمتغيرين (س،ص) مع إضافة خانة لحاصل مجموع (مجـ س ً × ص ً × ك) . أيضاً نقوم بعمل صف أفقى زائد إلى أعلى الجدول يتضمن قيم (س) من الجدول الهامشى للمتغير (س) ثم نضيف عمود أيضاً لقيم (ص) ملاصق لعمود فئات المتغير (ص) ولكن خارج الجدول. وبالمثل نحصل على قيم (ص) من الجدول الهامشى للمغير (ص).

٤- بدءاً من أول خلية على اليمين يتضمن التكرارات الأصلية في الجدول الثالث، نقوم بضرب جبري لقيمة صَ × سَ × قيمة التكرار في هذه الخلية (ك١) فنحصل على قيمة. أما بإشارة سالبة أو بإشارة موجبة فنسجل هذه القيمة في دائرة داخل نفس الخلية. وهكذا بالنسبة لباقي الخانات في الصف الأول، ويلاحظ في هذه الطريقة إننا نستخدم القيمة الأولى من (ص) في العمود الرأسي مع جميع قيم (س) في الصف الأفقي وفقاً لكل تكرار في هذا الصف. نقوم بعد ذلك بجمع القيم داخل الأقواس جبرياً ووضع صافي المجموع في العمود الأخير من الجول الدال على (مجـ سَ صَ ك). تكرار هذا العمل لكل الصفوف داخل الجدول، ففي الصف الثاني مثلاً نستخدم القيمة الثانية في الترتيب من قيم صَ في العمود الثالث مع جميع قيم سَ بواقع قيمة واحد من سَ لكل خلية، وبعد الانتهاء من جمع جميع الصفوف نقوم بإيجاد المجموع الكلي جبرياً (بإشاراته) فيكون هو قيمة المقدار مجـ سَ صَ ك.

مـثال ٣ :

أحسب معامل الارتباط (بيرسون) للعلاقة بين دخل عينة من الأسر (س)، ومتوسط إنفاقهم الشهري (ص) بالجنية المصري من البيانات المدونة بالجدول رقم (٦-٢).

جدول رقم (٦-٢)

ص \ س	١٥٠-	١٦٠-	١٧٠-	١٨٠-	١٩٠-	٢٠٠-	٢١٠-	المجموع
١٠٠-	٣	٢						٥
١١٠-	٢	٦	٢					١٠
١٢٠-	٣	٥	٣	٥	٢			١٨
١٣٠-		٥	٦	٤	١			١٦
١٤٠-	١	٣		٣	٥			١٢
١٥٠-		٢	٣		٤	١		١٠
١٦٠-	٥		٢			٤		١١
١٧٠-		١		٤	٢		٣	١٠
١٨٠-			١			٤	٣	٨
المجموع	١٤	٢٤	١٧	١٦	١٤	٩	٦	١٠٠

جدول التوزيع الهامشى للمتغير (ص)

ص	ك	مركز الفئة	صَ	صً	سً ك	صً ك
١٠٠ –	٥	١٠٥	– ٢٠	– ٢	– ١٠	٢٠
١١٠ –	١٠	١١٥	– ١٠	– ١	– ١٠	١٠
١٢٠ –	١٨	١٢٥	صفر	صفر	صفر	صفر
١٣٠ –	١٦	١٣٥	+ ١٠	+ ١	١٦	١٦
١٤٠ –	١٢	١٤٥	+ ٢٠	+ ٢	٢٤	٤٨
١٥٠ –	١٠	١٥٥	+ ٣٠	+ ٣	٣٠	٩٠
١٦٠ –	١١	١٦٥	+ ٤٠	+ ٤	٤٤	١٧٦
١٧٠ –	١٠	١٧٥	+ ٥٠	+ ٥	٥٠	٢٥٠
١٨٠ –	٨	١٨٥	+ ٦٠	+ ٦	٤٨	٢٨٨
المجموع	١٠٠				– ٢٠ + ٢١٣ ١٩٢	٨٩٨

∴ مجـ صً ك = ١٩٢

مجـ صً ك = ٨٩٨

جدول التوزيع الهامشى للمتغير (س):

س	ك	مركز الفئة	سَ	سً	سً ك	سً^{٢}ك
١٥٠–	١٤	١٥٥	١٠–	١–	١٤–	١٤
١٦٠–	٢٤	١٦٥	صفر	صفر	صفر	صفر
١٧٠–	١٧	١٧٥	١٠	١	١٧	١٧
١٨٠–	١٦	١٨٥	٢٠	٢	٣٢	٦٤
١٩٠–	١٤	١٩٥	٣٠	٣	٤٢	١٢٦
٢٠٠–	٩	٢٠٥	٤٠	٤	٣٦	١٤٤
٢١٠–	٦	٢١٥	٥٠	٥	٣٠	١٥٠
المجموع	١٠٠				١٥٧ – ١٤ ١٤٣	

مجـ سَ ك = ١٤٣

مجـ $سَ^٢$ ك= ٥١٥

* نقسم صَ على مدى الفئة لنحصل على صَ لتسهيل العمليات الحسابية ونفس الشيء بالنسبة للمتغير س.

(جـ) حساب مجـ سَ صَ ك

قيم صَ	قيم سَ	-١	صفر	١	٢	٣	٤	٥	مجـ سَ صَ ك
	ص \ س	١٥٠-	١٦٠-	١٧٠-	١٨٠	١٩٠-	٢٠٠	٢١٠-	المجموع
-٢	١٠٠ -	٣ (٦)	صفر						٦
-١	١١٠ -	٢ (٢)	٦ (صفر)	٢ (-٢)					صفر
صفر	١٢٠ -	٣	٥	٣	٥	٢			صفر
١	١٣٠ -		٥ (صفر)	٦ (٦)	٤ (٨)	١ (٣)			١٧
٢	١٤٠ -	١ (-٢)	٣ (صفر)		٣ (١٢)	٥ (٣٠)			٣٠
٣	١٥٠ -		٢ (صفر)	٣ (٩)		٤ (٣٦)	١ (١٢)		٥٧
٤	١٦٠ -	٥ (-٢٠)		٢ (٨)			٤ (٦٤)		٥٢
٥	١٧٠ -		١ (صفر)		٤ (٤٠)	٢ (٣٠)		٣ (٧٥)	١٤٥
٦	١٨٠ -			١ (٦)			٤ (٩٦)	٣ (٩٠)	١٩٢
	المجموع								٤٩٩

∴ مجـ س َ ص َ ك = ٤٩٨

وبتطبيق المعادلة رقم (٦-٤):

$$ر = \frac{\left(\frac{٤٩٩}{١٠٠}\right) - \left(\frac{١٤٣}{١٠٠}\right) \times \left(\frac{١٩٢}{١٠٠}\right)}{\sqrt{\left[\left(\frac{٥١٥}{١٠٠}\right) - \left(\frac{١٤٣}{١٠٠}\right)^{٢}\right] \times \left[\left(\frac{٨٩٨}{١٠٠}\right) - \left(\frac{١٩٢}{١٠٠}\right)^{٢}\right]}} = ٠,٥٥$$

معامل سبيرمان

يستخدم هذا المعامل فى حالة العينات صغيرة الحجم ويعتمد على ترتيب القيم فى كل متغير تحت الدراسة. فى نفس الوقت يعتبر معامل سبيرمان حالة خاصة من معامل ارتباط بيرسون لمتغيرين كلاهما يتم قياسه بالمقاييس الترتيبية. فلو فرضنا أن باحثاً يريد أن يعرف العلاقة بين حجم الفصل الدراسى للفرقةالنهائية فى اثنى عشر كلية جامعية لعام معين وليكن عام ١٩٩٥ وبين نسبة الخريجيين ممن يستكملون دراساتهم العليا للماجستير والدكتوراه وحصل على البيانات الموضحة بالجدول رقم (٦-٣):

الكلية	حجم الفصل الدراسى (س)	نسبة الدارسين (ص)
أ	٣٠٦٨	٢,٩
ب	٢٥٨٤	٣,٦
جـ	٢٠٦٧	١,٣
د	١٥٨٤	٦,٨
هـ	١٠٩٣	٤,٩
و	٨٤٧	١,٨
ز	٦٩٨	٤,٣
ح	٥٦٣	٨,٦
ط	٣٩٨	٥,٧

الكلية	حجم الفصل الدراسى (س)	نسبة الدارسين (ص)
ى	٣٠٤	٨,٩
ك	٢١٨	-٤,٧
ل	١٣٠	٧,٥

خطوات الحل :

١- نقوم بترتيب المتغير الأول (س) ترتيباً تصاعدياً أو تنازلياً ، فعلى سبيل المثال فى حالة الترتيب التصاعدى يتم إعطاء الرتبة الأولى لأقل درجة والرتبة الثانية للدرجة التى تليها وهكذا ويوضع ذلك فى عامود ترتيب الظاهرة (س) .

٢- نقوم بترتيب المتغير الثانى (ص) بنفس طريقة ترتيب المتغير (س) ويوضع فى العامود الثانى من نفس خانة ترتيب الظاهرة بالجدول.

٣- تقوم بحساب الفرق بين رتبة (س) ورتبة (ص) وذلك بطرح رتبة الثانى من رتبة الأول أو العكس. ويوضع الناتج من الخانة المسماة بالفرق بين الترتيبين ويرمز لفرق بالرمز (ف) .

٤- نقوم بعد ذلك بتربيع الفرق ($ف^٢$) ويوضع فى العامود الثانى المسماة الفرق بين الترتيبين ومربعه.

٥- نقوم بجمع القيم الموجودة فى العامود ($ف^٢$) لإيجاد مجـ $ف^٢$.

٦- يتم تطبيق معامل ارتباط الرتب لسبيرمان وصيغته كالآتى:

$$\text{معامل سبيرمان} = ١ - \frac{٦ \text{ مجـ } ف^٢}{ن (ن^٢ - ١)} \quad \text{معادلة رقم (٦ - ٥)}$$

حـل المـثال

الكلية	حجم الفصل (س)	نسبة الدارسين (ص)	ترتيب س	ترتيب ص	الفرق بين ف	مربع الفرق ف۲
أ	٣٠٦٨	٢,٩	١	١٠	٩ –	٨١
ب	٢٥٨٤	٣,٦	٢	٩	٧ –	٤٩
جـ	٢٠٦٧	١,٣	٣	١٢	٩ –	٨١
د	١٥٨٤	٦,٨	٤	٤	صفر	صفر
هـ	١٠٩٣	٤,٩	٥	٦	١ –	١
و	٨٤٧	١,٨	٦	١١	٥ –	٢٥
ز	٦٩٨	٤,٣	٧	٨	١–	١
ح	٥٦٣	٨,٦	٨	٢	٦	٣٦
ط	٣٩٨	٥,٧	٩	٥	٤	٣١
ى	٣٠٤	٨,٩	١٠	١	٩	٨١
ك	٢١٨	٤,٧	١١	٧	٤	١٦
ل	١٣٠	٧,٥	١٢	٣	٩	٨١
مجـ					صفر	٤٦٨

وباستخدام المعادلة رقم (٦-٥) يكون الحل كالآتى:

$$\text{معامل سبيرمان} = ١ - \frac{٦ \times ٤٦٨}{١٢ (١٤٤ - ١)} = ٠.٦٣٦$$

ويلاحظ أن المعادلة السابقة لسبيرمان قد تم استنتاجها من معادلة قيمة معامل الارتباط (ر) بشروط افتراضية:

١ - تساوى الوسطين الحسابيين للمتغيرين (س،ص).

٢ -تساوى الانحراف المعيارى للمتغيرين.

٣ - ترتيب القيم الأصلية ترتيباً تصاعدياً أو تنازلياً مع استبدال تلك القيم بأرقام مسلسلة يمكن وصفهما فى الجداول المفردة فقط. ومن ثم لا يصح استخدام معامل سبيرمان فى قياس الارتباط من الجداول التكرارية حتى لاتضيع خصائص القيم المفردة.

٤- من عيوب معامل ارتباط الرتب أنه إذا تغيرت القيم لن يتأثر قيم معامل الارتباط ، لكنه فى حالة معامل ارتباط بيرسون يكون عن طريق الانحرافات ومن ثم فأى تغيير فى القيم سوف يؤثر على قيم معامل الارتباط.

أما فى الحالات التى يجد الباحث فيها قيماً متشابهة لأى من المتغيرين (س،ص) فيقوم بإعطاء ترتيب متوسط للقيم المتشابهة كما يتضح ذلك من المثال التالى.

مثال رقم (٤):

أراد باحث اجتماعى أن يدرس العلاقة الارتباطية بين عدد الخريجين من قسم اللغة الفرنسية بكلية الآداب بجامعة عين شمس وفرص العمل المتاحة لهم فى مدينة القاهرة وذلك من خلال حصر عدد المشتغلين منهم خلال سبع سنوات متتالية بدءاً من عام ١٩٩٠ حتى عام ١٩٩٦ وأعطت الدراسة البيانات الموضحة فى الجدول رقم (٦-٤).

جـدول رقـم (٦-٤)

السنة	عدد الخريجين سنوياً (ص)	عدد المشتغلين بالقاهرة (س)
١٩٩٠	٧٤	١٠
١٩٩١	٦٥	٩
١٩٩٢	٧٩	١٢
١٩٩٣	٧٧	١٥
١٩٩٤	٦٩	١٢
١٩٩٥	٨٢	١٦
١٩٩٦	٨٦	١٧

تلاحظ من البيانات السابقة تساوى رقمى السنتين ١٩٩٢ ، ١٩٩٤ للمتغير (س) ومن ثم نأخذ المتوسط الحسابى للترتيب $\frac{٣ + ٤}{٢}$ = ٣,٥ فتأخذ كل منها فى الترتيب قيمة واحدة هى:

حـــل المـــثال:

	س	ص	رتبة س	رتبة ص	ف	ف^٢
١٩٩٠	١٠	٧٤	٢	٣	١-	١
١٩٩١	٩	٦٥	١	١	صفر	صفر
١٩٩٢	١٢	٧٩	٣,٥	٥	١,٥-	٢,٢٥
١٩٩٣	١٥	٧٧	٥	٤	١	١
١٩٩٤	١٢	٦٩	٣,٥	٢	١,٥-	٢,٢٥
١٩٩٥	١٦	٨٢	٦	٦	صفر	صفر
١٩٩٦	١٧	٨٦	٧	٧	صفر	صفر
مجـ					صفر	٦,٥

$$\text{معامل سبيرمان} = 1 - \frac{6 \times 6.5}{7\,(49 -)} = 0.88$$

معامل الارتباط فاى : Φ

وهو حالة خاصة من معامل بيرسون للارتباط ويستخدم عندما يكون المتغيران من النوع المتقطع المنقسم إلى قسمين نوعيين.

مثال رقم (٥) :

أراد باحث اجتماعى أن يعرف العلاقة بين النوع (ذكر وأنثى) لعينة من المبحوثين وانضمامهم لحزب سياسى معين وحيث أن المتغيرين من النوع المتقطع فمن الأجدر أن نستخدم تمييزاً رقمياً للنوع كأن يعطى للإناث رقم (١) وللذكور (صفر) ويكرر نفس العمل بالنسبة للحزبين السياسيين كأن يعطى للحزب الوطنى (١) الحزب المعارض (صفراً) والجدول رقم (٦-٥) يتضمن البيانات التى حصل عليها الباحث للمتغيرين س،ص.

جدول رقم (٦ – ٥)

المفردات	النوع	الحزب السياسى			
	س	ص	$س^2$	$ص^2$	س ص
أ	١	١	١	١	١
ب	١	١	١	١	١
جـ	١	صفر	١	صفر	صفر
د	١	١	١	١	١
هـ	١	١	١	١	١
و	صفر	صفر	صفر	صفر	صفر
ز	صفر	١	صفر	١	صفر
ح	صفر	١	صفر	١	صفر
ط	صفر	صفر	صفر	صفر	صفر
ى	صفر	صفر	صفر	صفر	صفر
مجـ	٥	٦	٥	٦	٤

وبتطبيق معادلة بيرسون للارتباط رقم (٦-٣) يكون حل المثال على النحو الآتى:

$$\text{ر} = \frac{\text{١٠ (٤) - (٥) (٦)}}{\text{[(١٠) (٥) - (٥)}^{\text{٢}}\text{] [(١٠) (٦) - (٦)}^{\text{٢}}\text{]}} \quad \text{٠,٤٠٩}$$

هذا المعامل يدل على وجود علاقة ارتباطية إيجابية منخفضة بين النوع والحزب السياسى الذى ينتمى إليه. وأن النتائج السابقة تشير إلى أن الإناث داخل تلك المجموعة من الأفراد تميل إلى الانتماء للحزب الوطنى بينما يميل الذكور للحزب المعارض. وهذا الاتجاه يعطى دلالة ارتباطية إيجابية.

أيضاً يمكن تنظيم البيانات بالجدول السابق فى جدول آخر يطلق عليه جدول (٢ × ٢) بحيث تحول الأرقام داخل كل مربع إلى تكرارت كما يتضح ذلك من الجدول (٦-٦).

جدول رقم (٦-٦)

الانتماء للأحزاب السياسية	ذكور	إناث	جملة
وطنى	٢	٤	٦
معارض	٣	١	٤
جملة	٥	٥	١٠

كذلك يمكن تعديل مكونات هذا الجدول إذا استخدمنا الرموز الرقمية (صفر،١) ثم نعطى كل خلية من الخلايا الأربعة حرفاً أبجدياً مثل أ ، ب ، جـ ، د حتى يمكن استخدام الرموز والأحرف فى المعادلة السابقة والجدول المعدل يتضح فى الجدول رقم (٦-٧).

جدول رقم (٦-٧)

صفر	١	جملة
أ		أ + ب
جـ	د	جـ + د
أ + جـ	ب + د	ن

يمكن تبسيط معادلة بيرسون السابقة لحساب معامل الارتباط (ر) صياغة أخرى تحسب قيمة معامل فاى على النحو التالي:

معادلة رقم (٦-٦)

$$\Phi = \frac{\text{ب جـ} - \text{أ د}}{\sqrt{(\text{أ} + \text{ب})\ (\text{جـ} + \text{د})\ (\text{أ} + \text{جـ})\ (\text{ب} + \text{د})}}$$

وباستخدام هذه المعادلة في المثال السابق، نحصل على نفس القيمة السابقة تقريباً لمعامل ارتباط بيرسون حيث:

$$= \frac{(٤)\ (٣) - (٢)\ (١)}{\sqrt{(٢ + ٤)\ (٣ + ١)\ (٢ + ٣)\ (٤ + ١)}} = ٠,٤٠٨$$

معامل التوافق C.

من خلال مناقشتنا لمعامل الارتباط فاى قلنا أن هناك بعض القيود التى تحد من استخدام هذا المعامل. فإن كان معاملى التوافق وفاى متشابهين إلى حد كبير من حيث أهميتهما فى قياس العلاقة بين متغيرين ، يتم قياسهما بواسطة مقياس تصنيفى ، فإن معامل فاى يصلح للبيانات الممكن تصنيفها فى جدول مزدوج أو ما يسمى بجدول (٢ × ٢)، أما ما يزيد على ذلك فلا يقدر معامل فاى على قياسه ومن ثم استطاع كارل بيرسون أن يصور معاملاً آخر أسماه معامل التوافق حيث يمكن استخدامه فى جداول أكبر فى تقسيماتها النوعية عن (٢ × ٢) مثال ذلك جداول (٣ ×٤) أو (٣ × ٥) وهكذا. ومعنى جدول (٣ × ٥) أن يتضمن ثلاث تصنيفات للمتغير الأول وخمس تصنيفات للمتغير الثانى.

مثال رقم (٦) :

يتضمن الجدول رقم (٦-٨) بيانات العلاقة بين المستوى التعليمى وعوامل التغيب عن العمل لمائة عامل فى إحد المصانع بشبرا الخيمة وأراد باحث أن يحسب العلاقة الارتباطية بين المتغيرين.

جدول رقم (٦-٨)

العلاقة بين المستوى التعليمى وأسباب التغيب عن العمل

الحالة التعليمية	مرض	مشكلات أسرية	مشاكل عمل	مجـ
تعليم جامعى	٥	١٠	٣٥	٥٠
تعليم متوسط	٦	٩	١٥	٣٠
أمى	٨	٧	٥	٢٠
مجـ	١٩	٢٦	٥٥	١٠٠

خطـوات حـل المثـال ·

١ - تربيع كل تكرار من التكرارت المدونة فى خلايا الجدول

٢ - نقسم مربع كل تكرار حصلنا عليه على حاصل ضرب مجموع تكرارت الصف فى مجموع تكرارت العمود الواقعة فى خانة هذا التكرار وذلك على النحو التالى:

$$\frac{\text{مربع تكرار الخلية}}{\text{مجموع تكرار العمود} \times \text{مجموع تكرار الصف}}$$

٣ - جمع خوارج القسمة للحصول على (مجـ).

٤- حساب معامل التوافق (C) باستخدام المعادلة الآتية:

$$C = \sqrt{1 - \frac{1}{\text{مجـ}}}$$

معادلة رقم (٦ - ٧)

وباستخدام المعادلة السابقة رقم (٦-٧) يكون معامل التوافق كالآتى:

حـل المـثال :

$$\text{مجـ الصف الأول} = \frac{(٥)^٢}{١٩ \times ٥٠} + \frac{(١٠)^٢}{٢٦ \times ٥٠} + \frac{(٣٥)^٢}{٥٥ \times ٥٠}$$

$$= \frac{٢٥}{٩٥٠} + \frac{١٠٠}{١٣٠٠} + \frac{١٢٢٥}{٢٧٥٠}$$

$$= ٠,٠٢٦ + ٠,٠٧٧ + ٠,٤٤٥ = ٠,٥٤٨$$

$$\text{مجـ الصف الثاني} = \frac{(٦)^٢}{١٩ \times ٣٠} + \frac{(١٩)^٢}{٢٦ \times ٣٠} + \frac{(١٥)^٢}{٥٥ \times ٣٠}$$

$$= \frac{٣٦}{٥٧٠} + \frac{٨١}{٧٨٠} + \frac{٢٢٥}{١٦٥٠}$$

$$= ٠,٠٦٣ + ٠,١٠٤ + ٠,١٣٦ = ٠,٣٠٣$$

$$\text{مجـ الصف الثالث} = \frac{(٨)^٢}{١٩ \times ٢٠} + \frac{(٧)^٢}{٢٦ \times ٢٠} + \frac{(٥)^٢}{٥٥ \times ٢٠}$$

$$= \frac{٦٤}{٣٨٠} + \frac{٤٩}{٥٢٠} + \frac{٢٥}{١١٠٠}$$

$$= ٠,١٦٨ + ٠,٠٩٤ + ٠,٠٢٣ = ٠,٢٨٥$$

$$\text{مجموع الصفوف} = ٠,٥٤٨ + ٠,٣٠٣ + ٠,٢٨٥ = ١,١٣٦$$

وباستخدام المعادلة السابقة رقم (٦ – ٧) يكون معامل التوافق كالآتى :

$$C = \sqrt{١ - \frac{١}{١,١٣٦}} = ٠,٣٤٦$$

وهناك صياغة أخرى لمعامل التوافق باستخدام (كا٢)على النحو التالى :

$$C = \sqrt{\frac{كا^٢}{كا^٢ + ن}} \qquad \text{معادلة (٦ – ٨)}$$

$$كا^2 = مجـ \left[\frac{(ك - ك')^2}{ك} \right]$$

ك = التكرارت التجريبية.

ك' = التكرارت المتوقعة.

ن = حجم العينة.

ملاحظات عامة على معامل التوافق:

١ – دائماً تكون قيمة هذا المعامل موجبة لأن قيمة $كا^2 \geqslant$ صفر .

٢ – لا يصلح معامل التوافق للمقارنة بين جدولين إلا بشرط واحد هو أن تتساوى أعداد الصفوف والأعمدة بينهما.

٣ – يستخدم لقياس ارتباط بين الصفات وبين متغيرات مقاسة وصفياً وهذا الشرط لا يتحقق باستخدام معامل الارتباط الخطى البسيط أو معامل ارتباط الرتب.

الارتباط الجزئى والمتعدد:

تناولنا فيما سبق العلاقة بين متغيرين إحداهما معتمد والآخر متغير مستقل. إلا أن تحليلات الارتباط والانحدار قد تمتد لتشمل أكثر من مقياس بحيث يكون إحداهما متغير معتمداً بينما الباقون متغيرات مستقلة. ومن ثم تصبح المشكلة محصورة فى توقعات قيم المتغير المعتمد (ص) المناظرة لقيم المتغيرات المستقلة $س_1$، $س_2$ ، $س_3$، ...... ، $س_ن$ ولتسهيل تلك العملية على القارئ فسوف نناقش فيما بعد النموذج المبسط من الانحدار الخطى.

أشرنا فيما سبق إلى الأنواع المختلفة للارتباط، فإلى جانب الارتباط الخطى البسيط يوجد الارتباط الجزئى الذى يستخدم لقياس العلاقة بين متغيرين مع افتراض ثبوت التأثيرات التى قد تحدثها متغيرات أخرى أو ما يسميه الباحثون بتثبيت المتغيرات. والنوع الثالث هو الارتباط المتعدد والذى يستخدم للدلالة على مدى وحجم التباين الكلى الحادث فى المتغير المعتمد والذى يمكن تفسيره من خلال تفاعل المتغيرات المستقلة فيما بينها. وسوف نناقش كيفية حساب معامل

الارتباط الجزىء من خلال خطوات الحل للمثال التالى.

مثال رقم (٧) :

أعطت نتائج إحدى المسوح الاجتماعية التى أجريت فى أحدى المدن الامريكية أن معامل الارتباط بين التمييز فى الأجور والتفرقة العنصرية كانت قيمته بالنسبة للزنوج (٠,٥٣٦) . أيضاً عندما قيس معامل الارتباط بين المتغير الأول وتوزيع السكان (ريف/حضر) كان قيمة معامل الارتباط بالنسبة لسكان الحضر (٠,١٣٩) . بينما كانت العلاقة الارتباطية بين المتغيرين الثانى والثالث (−٠,٢٤٨) وإذا أردنا من خلال معاملات الارتباط السابقة أن نعرف مدى العلاقة بين المتغير الأول وهو التمييز فى الأجور مع المتغير الثالث وهو نوعية السكان مع تثبيت تأثير المتغير الثانى فإننا نستخدم المعادلة الآتية لقياس معامل الارتباط الجزئى حيث أن:

$ر_{٢.٣١}$ = معامل الارتباط الجزئى.

$ر_{٣١}$ = معامل الارتباط بين المتغيرين الأول والثالث.

$ر_{٢١}$ = معامل الارتباط بين المتغيرين الأول والثانى.

$ر_{٣٢}$ = معامل الارتباط بين المتغيرين الثانى والثالث.

$$ر_{٢.٣١} = \frac{(ر_{٣١}) - (ر_{٢١})\,(ر_{٣٢})}{\sqrt{(١ - ر_{٢١}^{٢}) \times (١ - ر_{٣٢}^{٢})}}$$

معادلة رقم (٧ – ١١)

الانحـــدار الخطـــى

إحصائياً، يرجع استخدام لفظ « انحدار » إلى عام ١٨٨٥ عندما استخدمها فرنسيز جالتون Galton فى مقالة الذى نشره خلال ذلك العام والذى ضمنه نتائج دراسته على العلاقة بين أطوال الآباء وأبنائهم وأن هناك انحدارا لطول الأبناء نحو متوسط أطوال المجتمع الأصلى تحت الدراسة . كما خلص إلى نتيجة هامة حين ذكر أن قيم أطوال الأبناء تنحدر نحو موضع ما يقع ما بين أطوال آبائهم والقيمة المتوسطة (المتوسط) للمجتمع الأصلى. ولقد استفاد بهذه النتيجة كارل بيرسون فيما بعد بما أسماه معامل الانحدار. فإذا كان معامل الارتباط. كما يرى بيرسون- يعطى تلخيصاً واضحاً للعلاقة بين متغيرين س،ص فإن معامل الانحدار يعبر عن التغير المتوقع فى المتغير (ص) (بوصفه متغيراً تابعاً) كلما تغيرت قيم المتغير (س) المناظرة على أساس أنه متغير مستقل. ومن ثم يتحدد الهدف الأساسى لمعامل الانحدار فى قياس تأثير المتغير (س) على المتغير التابع (ص) العلاقة فى شكل معادلة إنحنائية أو خطية، وما يهمنا فى هذا الجزء هو الانحدار الخطى واستخدام معادلة تفسر هذا النوع من الانحدار والتى يمكن صياغتها على الصورة التالية:

ص = أ + ب س

حيث أ = مقدار ثابت يساوى قيمة المتغير (ص)إذا كانت قيمة (س) تساوى صفراً. فى المعادلة الموضحة. وتقاس قيمة (أ) على المحور الصادى وذلك فى حالة انحدار (ص) على (س).

ب = الميل Slope لخط الانحدار على المحور الأفقى والذى يساوى جبرياً ظل زاوية ميل خط الانحدار على المحور الأفقى. كما يمثل أيضاً كمية التغير فى قيمة المتغير (ص) المصاحبة لكل وحدة تغير من وحدات تغير المتغير المستقل (س).

ويمكن حساب قيمتى كل من (أ) ، (ب) فى المعادلة السابقة من واقع البيانات التجريبية على النحو الذى سنعرض له فيما بعد.

أهم الطرق الشائعة فى دراسة الانحدار للبيانات غير المبوبة :

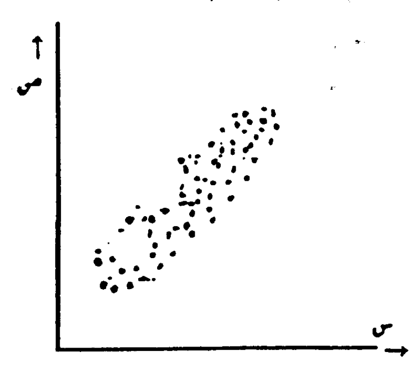
١- الشكل الانتشارى:

ويستخدم للتعرف مبدئياً على شكل العلاقة بين المتغيرين (س،ص) وذلك باستخدام محاور الإحداثيات (المحور السينى والمحور الصادى) حيث يتم رصد وتمثيل كل قيم المتغير الأول مع ما يناظرها من قيم المتغير الثانى فى نقطة توقع على الشكل حيث يكون لكل نقطة قيمتين (س،ص) يحددان موضعها، ونستمر فى رصد جميع النقاط حتى نحصل على شكل انتشارى لجميع قيم (س) ، (ص)، ويمكن من الشكل الانتشارى تحديد ماهية العلاقة ، وهل توجد أم تنعدم بين المتغيرين ، هذا فضلاً عن معرفة إتجاه تلك العلاقة فى حالة وجودها.

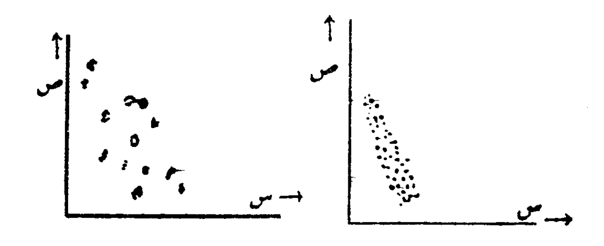
ففى الشكل رقم (٦-١٢) نجد شكلاً انتشارياً لقيم (س،ص) يغلب عليها الاتجاه ناحية اليمين ابتداءاً من ناحية نقطة الأصل. كما يلاحظ وجود نوع من التجانس فى القيم أى تقل خاصية التشتت. ويطلق على هذا الشكل بالانتشار الموجب.

وفى شكل رقم (٦-١٣) يتضح أيضاً وجود تجانس إلى حد ما بين القيم وانخفاض تشتتها وتنافرها وأن اتجاه الانتشار ناجية اليسار. ومن ثم يطلق على هذا الشكل بالشكل الانتشارى السالب.

أما فى الشكل رقم (٦-١٤) فنلاحظ عدم انتظام النقط وتشتتها على الشكل الانتشارى بحيث يصعب رسم خط مستقيم يربط بين معظم تلك النقط ، ومن ثم لا توجد أى علاقة بين المتغيرين (س،ص) .



شكل رقم (٦ - ١٢) توزيع انتشارى موجب



شكل رقم (٦ – ١٣)
توزيع انتشارى سالب

شكل رقم (٦-١٤) توزيع انتشارى يوضح
عدم وجود علاقة بين المتغيرين (س ، ص)

مثال رقم (٨) :

ارسم الشكل الانتشارى للعلاقة بين المتغيرين (س،ص) من واقع البيانات الموضحة بالجدول رقم (٦-٩).

جدول رقم (٦-٩)

مسلسل	س	ص	س^2	س ص
١	١٥	١٢	٢٢٥	١٨٠
٢	١٠	١٣	١٠٠	١٣٠
٣	٧	٩	٤٩	٦٣
٤	١٨	١٨	٣٢٤	٣٢٤
٥	٥	٧	٢٥	٣٥
٦	١٠	٩	١٠٠	٩٠
٧	٧	١٤	٤٩	٩٨
٨	١٧	١٦	٢٨٩	٢٧٢
٩	١٥	١٠	٢٢٥	١٥٠
١٠	٩	١٢	٨١	١٠٨
١١	٨	٧	٦٤	٥٦
١٢	١٥	١٣	٢٢٥	١٩٥

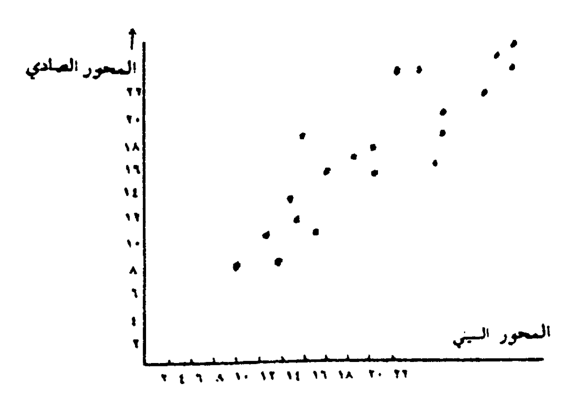
مسلسل	س	ص	$س^2$	س ص
١٣	١١	١٤	١٢١	١٥٤
١٤	١٧	١٩	٢٨٩	٣٢٣
١٥	٨	١٠	٦٤	٨٠
١٦	١١	١٦	١٢١	١٧٦
١٧	١٢	١٢	١٤٤	١٤٤
١٨	١٣	١٦	١٦٩	٢٠٨
١٩	١٨	١٩	٣٢٤	٣٤٢
٢٠	٧	١١	٤٩	٧٧

الحـــل:

نرسم المحورين (س،ص) ثم نوقع عليها كل قيمة للمتغير (س) وما يناظرها فى الجدول من قيمة للمتغير (ص) ونكرر ذلك العمل فى العشرين حالة المعطاة فنحصل على عشرين نقطة منتشرة كما هو مبين بالشكل الانتشارى رقم (٦-١٥).

وكمثال النقطة الأولى إحداثياتها (س،ص) هى (١٢،١٥) فنأخذ بمقياس رسم مناسب قيمة (١٥) على المحور السينى ابتداء من نقطة الأصل وعند القيمة نقيم خطاً رأسياً موازياً للمحور الصادى. وبعد ذلك نأخذ مقياس رسم مناسب على المحور الصادى ونرصد قيمة (١٢) على هذا المحور فتتحدد ، ومنها نرسم خطاً أفقياً فى إتجاه المحور السينى وموازى له فيلتقى الخطان الرأسى والأفقى عند نقطة تمثل الحالة الأولى فى خانة المسلسل بالجدول. ونكرر العمل بالنسبة لباقى الحالات حتى نصل إلى الحالة العشرين والأخيرة.

ولتحديد خط الإنحدار يجب أن نختار خطاً يتوسط جميع النقاط فى الشكل الانتشارى السابق وهناك طريقتان لعمل ذلك، أما أن نقوم بواسطة اليد رسم هذا الخط ولهذه الطريقة عيوبها حيث يعتمد رسم الخط على مهارة الدارس، وأما أن نستخدم طريقة رياضية وهى طريقة المربعات الصغرى.



شكل رقم (٧ – ١٥) الشكل الانتشارى للعلاقة بين س ، ص

٢- طريقة المربعات الصغرى:

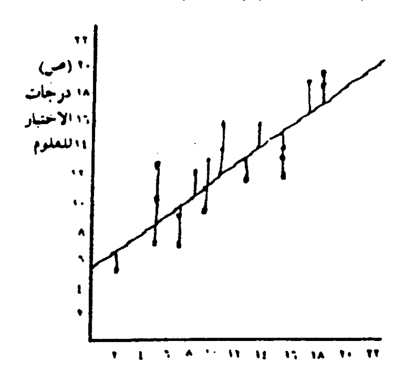
لتلخيص ودقة إبراز العلاقة بين المتغيرين (س،ص) ، تستخدم طريقة المربعات الصغرى ويمكن باستخدام تلك الطريقة تمثيل العلاقة بين المتغيرين (س،ص) بخط مستقيم يمر خلال النقط فى الشكل الانتشارى وأن أفضل خط مستقيم يمثل الإنحدار هو ذلك الذى يمر بمعظم القيم المركزية أو يمر بالمسار المركزى عبر النقط فى الشكل الانتشارى ويعرف المسار المركزى بأنه الخط الذى تكون قيمة مجموع مربع المسافات حوله بين النقط أقل ما يمكن . وهذا الخط أو المسار المركزى يعتبر خط الإنحدار المنشود.

وإحصائياً يمكن القول أن خط الإنحدار خط متوسط يعبر عن القيم المتناظرة للمتغيرين (س،ص) بحيث أن مجموع إنحرافات قيم (ص) الفعلية عن قيم المتوسط الحسابى للمتغير (ص) يساوى صفراً. ويمكن أن يلحظ الدارس أن من خصائص المتوسط الحسابى ، كما ذكرنا سابقاً أن تكون قيمة مجموع مربع إنحرافات القيم الفعلية عن المتوسط أقل ما يمكن.

ونخلص مما سبق، أن خط الإنحدار للشكل الانتشارى يعتبر نقطة الإتزان للتوزيع الثنائى المرتبط فضلاً عن فائدته للتنبؤ عن قيم المتغير التابع (ص) فى

(ص) فى علاقته بالمتغير المستقل (س).

وفى المثال السابق لو قمنا بتوصيل خطوط رأسية بين النقط على جانبى خط الإنحدار نجدها قريبة جداً من هذا الخط وبشكل أقرب للانتظام منه للانتشار والتفرق كما يتضح ذلك من الشكل رقم (٦-١٦).



(س) درجات الاختبار (للرياضة)

شكل رقم (٦-١٦) خط الانحدار باستخدام خاصية المربعات الصغرى

معادلة إنحدار ص على س

قلنا أن طريقة المربعات الصغرى تعطى أكثر الخطوط توفيقاً لإنحدار المتغير الأول وليكن (ص) على المتغير الثانى (س). وأن معادلة هذا الخط تكون على الصورة.

ص = أ + ب س   معادلة رقم (٦-١٢)

وتسمى بمعادلة خط الإنحدار (ص على س) حيث (أ) هى الجزء المقطوع intercept من المحور الصادى، (ب) هى ميل خط الإنحدار.

مثال رقم (٩):

استخدم بيانات جدول رقم (٦-١)لإيجاد معادلة الإنحدار الخطى ثم أرسم خط إنحدار (ص) على (س) أو الرضا عن العمل على الأجر اليومى للأثنى عشر عاملاً.

الحـــــل :

المتغير المستقل هو الأجر اليومى (س) .

المتغير التابع هو الرضا عن العمل (ص) .

من المعادلة ص = أ + ب س نوجد قيمتى (ب) ، (أ) من المعادلتين التاليتين.

$$\text{ب} = \frac{\text{ن مجـ س ص} - (\text{مجـ س})\ (\text{مجـ ص})}{\text{ن مجـ س}^{٢} - (\text{مجـ س})^{٢}}$$

معادلة رقم (٦ – ١٣)

$$\text{أ} = \frac{\text{مجـ ص} - \text{ب مجـ (س)}}{\text{ن}}$$

معادلة رقم (٦ – ١٤)

وللحصول على قيم (مجـ س ص)، (مجـ س) ، (مجـ ص) ، (مجـ س^{٢})

نقوم بعمل الجدول رقم (٦ –١٠) .

جدول رقم (٦–١٠)

العمال	الدخل س	الرضا عن العمل ص	س^{٢}	س ص
١	١٠,٥٠	٩٤	١١٠,٢٥	٩٨٧,٠٠
٢	٩,٥٠	٨٩	٩٠,٢٥	٧٤٥,٥٠
٣	٩,٠٠	٩١	٨١,٠٠	٨١٩,٠٠
٤	٨,٢٥	٩٠	٦٨,٠٦	٧٤٢,٥٠
٥	٨,٠٠	٨٤	٦٤,٠٠	٦٧٢,٠٠
٦	٧,٥٠	٩٢	٥٦,٢٥	٦٩,٠٠
٧	٦,٢٥	٨٦	٣٩,٠٦	٥٣٧,٥٠
٨	٦,٠٠	٨١	٣٦,٠٠	٤٨٦,٠٠
٩	٥,٧٥	٨٦	٣٣,٠٦	٤٩٤,٥٠
١٠	٥,٥٠	٨٢	٣٠,٢٥	٤٥١,٠٠
١١	٤,٥٠	٧٤	٢٠,٢٥	٣٣٣,٠٠
١٢	٤,٢٥	٨١	١٨,٠٦	٣٤٤,٢٥
	٨٥,٠٠	١٠٣٠	٦٤٦,٤٩	٧٤٠٢,٢٥

$$ب = \frac{١٢ (٧٤٠٢,٢٥) - (٨٥) (١٠٣٠)}{١٢ (٦٤٦,٤٩) - (٨٥)^٢}$$

$$= \frac{٨٨٨٢٧ - ٨٧٥٥٠}{٧٧٥٧,٨٨ - ٧٢٢٥}$$

$$= \frac{١٢٧٧}{٥٣٢,٨٨} = ٢,٣٩٦$$

$$أ = \frac{١٠٣٠ - (٢,٣٩٦) (٨٥)}{١٢}$$

$$= \frac{١٠٣٠ - ٢٠٣,٦٦}{١٢}$$

$$= \frac{٨٢٦,٣٤}{١٢} = ٦٨,٨٦$$

وممكن استخدام القيمتين (أ) (ب) فى رسم خط الإنحدار بأن نبدأ بإيجاد قيمة (ص) عند (س) = صفر من المعادلة رقم (٦ - ١٢)

∴ ص = ٦٨,٨٦ + ٢,٣٩٦ (صفر)

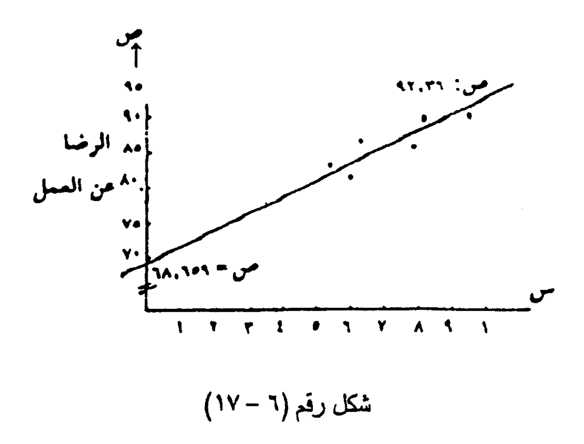
= ٦٨,٨٦

ثم أوجد قيمة ص على س عندما تكون س = ١٠

∴ ص = ٦٨,٨٦ + ٢,٣٩٦ × ١٠

= ٩٢,٨٢

وهكذا نرسم خط الإنحدار مع ملاحظة أن قيمة ص = ٩٢,٨٢ سوف تقع على هذا الخط. وسوف نحصل على خط الإنحدار كما يصوره الشكل رقم (٦-١٧).



شكل رقم (٦ – ١٧)

تقدير الخطأ المعيارى :

فى المثال السابق، أمكن رسم خط الإنحدار كما هو موضح بالشكل رقم (٦-١٧) . إلا أن اختيار أفضل خط للإنحدار يتطلب أن تحسب مقدار الخطأ وتصحيح القيم للمتغير التابع (ص) . ولتوضيح ذلك نأخذ على سبيل المثال من جدول الأجور والرضا عن العمل قيمة س = ٦ وعندها نجد أن قيمة ص = ٨١ ولكن لا يمكن الجزم بأنها هى القيمة التنبؤية والدليل على ذلك أن نعوض فى المعادلة بقيمتى أ، ب ص = ٦ فنحصل على:

ص = أ + ب س

= ٦٨,٨٦ + ٢,٣٩٦ (٦) = ٨٣،٢٤

ومن ثم فإن قيمة (ص) الفعلية فى الجدول بمقدار ٨٣,٢٤ – ٨١ = ٢٤ ، ٢ وهذا الفرق يمثل مقدار الخطأ بين القيمة المتوقعة والفعلية. ويمكن القول أن الخطأ المرتبط بخط الإنحدار يمكن مقارنته بالخطأ فى استعمال المتوسط كقيمة تقديرية لتوزيع جميع القيم. ولما كان خصائص المتوسط – كما قلنا سابقاً – أن مجموع الانحرافات حول المتوسط تساوى صفراً، فإن خط الإنحدار يأخذ هذه الخاصية

بحيث أن الفارق بين مسافات القيم المنتشرة قريباً لأعلى أو لأسفل من هذا الخط تساوى صفراً أيضاً . ومن ثم يمكن إعادة تقدير قيمة الخطأ لكل الحالات فى جدول خاص حتى يمكن التوصل إلى أفضل خط يمثل الإنحدار.

العينةص	قيمة ص المقدرة بالمعادلة ص= أ + ب س	الفرق بين قيمتى ص الفعلية والمقدرة
٩٤	٩٤,٠٢	– ٥,٠٢
٨٩	٩١,٦٢	– ٢,٦٢
٩١	٩٠,٤٢	٠,٥٨
٩٠	٨٨,٦٣	١,٣٧
٨٤	٨٨,٠٣	– ٤,٠٣
٩٢	٨٦,٨٣	٥,١٧
٨٦	٨٣,٨٤	٢,١٦
٨١	٩٣,٢٤	٢,٢٤–
٨٦	٨٢,٦٤	٣,٣٦
٨٢	٨٢,٠٤	– ٠,٠٤
٧٤	٧٩,٦٤	– ٥,٦٤
٨١	٧٩,٠٤	١,٩٦

ونلاحظ أن الفرق بين القيمتين ≠ الصفر نظراً لأن هناك تقريباً فى حساب الخطأ .

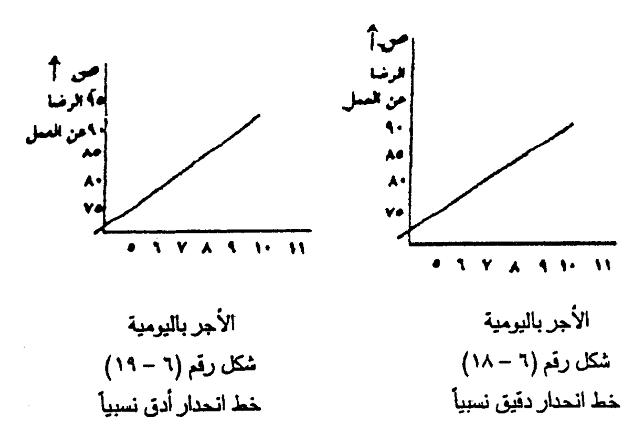
ويحسب الخطأ المعيارى بطريقتين أولهما هذه الطريقة بحساب قيمة ص المقدرة فى الجدول السابق وباستخدام المعادلة.

$$\text{الخطأ المعيارى} = \sqrt{\frac{\text{مجـ}\,(\text{ص} - \text{ص}_{\text{مقدرة}})^{2}}{\text{ن} - ٢}}$$

وأيضاً يمكن تفادى حساب ص المقدرة وحساب الخطأ المعيارى بدلالة قيم أ، ب فى معادلة الانحدار الخطى وذلك من المعادلة التالية:

$$\text{الخطأ المعيارى} = \sqrt{\frac{\text{مجـ ص}^{2} - \text{أ مجـ ص} - \text{ب مجـ س ص}}{\text{ن} - ٢}}$$

ومن ثم على أساس تقدير قيمة الخطأ المعيارى يمكن رسم خط الإنحدار البسيط بدقة أكثر حيث أن حساب القيم المتوقعة للمتغير (ص) ورصد ذلك على الشكل الانتشارى فى المثال السابق للعلاقة بين الأجور والرضا عن العمل، سوف نجد أن معظم النقط تقترب أكثر من خط الانحدار بدلاً من تناثرها بعيداً وتشتتها عن هذا الخط كما يتضح ذلك من الشكلين رقم (٦-١٨) ، (٦-١٩) لأفضل خطى إنحدار للمثال السابق.



الأجر باليومية
شكل رقم (٦ – ١٨)
خط انحدار دقيق نسبياً

الأجر باليومية
شكل رقم (٦ – ١٩)
خط انحدار أدق نسبياً

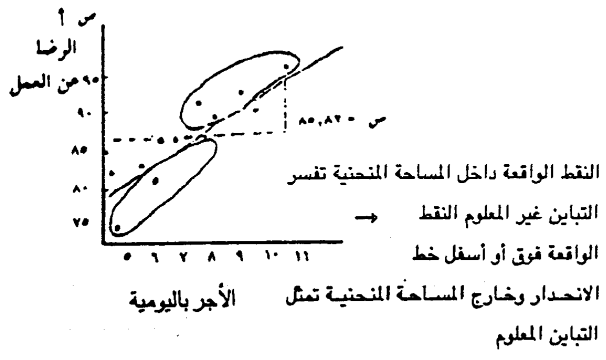
مربع قيمة معامل ارتباط ر٢

يشير ر٢ إلى التباين الحادث فى المتغير التابع وهو فى المثال السابق، الرضا عن العمل وذلك تبعاً للمتغير المستقل وهو الأجر. ولو طرحنا قيمة ر٢ من الواحد الصحيح فإن ناتج الطرح يمثل قدراً من التباين الذين يرجع إلى أسباب غير معلومة ومقدراه فى هذا المثال كالآتى:

ر = ٠,٨٣٧٤ ∴ ر٢ = ٠,٧٠١٢

مقدار حجم التباين غير المعلوم = ١-٠,٧٠١٢ = ٠,٢٩٨٨

وهذا يعنى أن ٢٩,٨٨ ٪ من التباين فى كل من الأجر بالساعة والرضا عن العمل يرجع إلى بعض المتغيرات الأخرى ويمكن باستخدام الشكل الانتشارى توضيح جزئى التباين المعلوم وغير المعلوم فكما يتضح فى الشكل رقم (٦-٢٠) نجد أن المشاهدات الواقعة بين ص = ٨٥,٨٣ وخط الإنحدار تمثل قيم التباين المفسرة بينما المشاهدات التى تقع تحت أو أسفل خط الإنحدار فإنها تمثل التباين غير المعلوم لأنها تتجاوز قيم ص المتوقعة.



شكل رقم (٦-٢٠) الشكل الانتشارى بين القيمة المعلومة وأيضاً غير المعلومة للتباين بدلالة (ر٢)

١- ارتباط خطى بسيط:

يقاس بمعامل الارتباط وقيمته تتراوح بين (+١) ،(-١) . وأساس حسابه هو استخدام إنحرافات كل من المتغيرين المعتمد والمستقل عن المتوسط الحسابى لكل منهما. وهى طريقة بيرسون فى حساب معامل الارتباط.

٢- الارتباط الجزئى:

هو العلاقة بين ظاهرتين مع إبقاء العوامل الأخرى ثابتة.

٣- الارتباط المتعدد:

يشير إلى العلاقة الارتباطية بين متغير تابع ومتغيرات أخرى مستقلة.

٤- معامل سبيرمان:

يعتبر حالة خاصة من معامل ارتباط بيرسون (ر)، إلا أنه يستخدم فى حالة العينات صغيرة الحجم ويقيس العلاقة الارتباطية بين متغيرين يتم قياسهما بالمقاييس الترتيبية.

٥- معامل الارتباط فاى:

وهو حالة خاصة من معامل بيرسون للارتباط ويستخدم عندما يكون المتغيران (س،ص) من النوع الوصفى المنقسم إلى قسمين نوعيين أى يمكن تصنيفهما فى جدول مزدوج (٢ × ٢) .

٦- معامل التوافق:

يستخدم فى حالة المتغيرات التى يتم قياسها وصفياً وتصنف فى جداول ذات تقسيمات نوعية أكبر من (٢ × ٢) مثال ذلك جداول (٣ × ٤) أو (٣ × ٥) ....إلخ.

طريقة المربعات الصغرى:

وهذه الطريقة تسمى بالمربعات الصغرى لأنها تجعل مجموع مربعات الأبعاد الرأسية لجميع النقاط عن الخط المستقيم الذى نريدَ توفيقه أقل ما يمكن ومعادلة هذا الخط تكون على الصورة التالية:

ص = أ + ب س

* * * * *

تمـــــارين

١ - توضح البيانات الآتية درجات عشرة من طلاب قسم الاجتماع فى مادتى الإحصاء الاجتماعى (س) ومناهج البحث (ص) المطلوب حساب معامل بيرسون للارتباط ومدى دلالته الإحصائية.

س	١٨	١٤	١٢	٨	١٢	١١	١٦	١٥	١٦	٨
ص	٩	١٥	١١	٧	١١	٩	١٣	١٦	١٧	٧

٢- باستخدام القيم الخام الآتية للمتغيرين (س)، (ص) أحسب معامل بيرسون للارتباط ثم استخدم الوسط الفرضى لكل من المتغيرين.

س	٦٤	٧١	٧٠	٦٨	٧٠	٦٨	٧٣	٦٩	٧٠	٦٧
ص	١٥٦	١٧٨	١٧٥	١٦٧	١٦٤	١٦٠	١٩٨	١٦٩	١٨٣	١٨٠

٣ - فيما يلى تقديرات عينة من الطلبة فى امتحان مادتى الإحصاء والرياضة والمطلوب حساب معامل سبيرمان بين تقديرات المادتين.

رقم الطالب	١	٢	٣	٤	٥	٦
الإحصاء	ض	م	حـ	ضحـ	ل	حـحـ
الرياضة	ل	حـحـ	حـ	ض	ض	م

٤ - أمكن التوصل إلى البيانات الآتية عن المتغيرين (س) ، (ص)

مجـ س = ٤٨   مجـ ص = ٧٦

مجـ س٢ = ٥٣٦   مجـ ص٢ = ١١٠٨

مجـ س ص = ٧٢٠   ن = ٦

والمطلوب حساب معامل بيرسون للارتباط بين قيم س،ص.

٥ - الجدول الآتى يوضح توزيع إعمار كل من الأزواج (س) والزوجات (ص) فى عدد من الأسر والمطلوب قياس الارتباط بينهما.

ص	٢٠-	٣٠-	٤٠-	٥٠-	٦٠مجـ
١٥-	٤	١٦	-	-	٢٠
٢٥-	٦	١٤	٥	-	٢٥
٣٥- ٤٥	-	-	٤	١	٥
مجـ	١٠	٣٠	٩	١	٥٠

٦- باستخدام المعادلة العامة لخط الانحدار (ص = أ + ب س) أحسب قيمة (ص) من البيانات التالية:

(أ) س = ١٢ ، أ = ٥ ، ب = ٣

(ب) س = ٣,٥ ، أ = ٢,٧ ، ب = ٨,١

(جـ) س = ١٤ ، أ = ٩ ، ب = ٤

(د) س = ١٠ ، أ = ٤,٥ ، ب = - ٣

(هـ) س = ١١,٢ ، أ = ٣ ، ب = ب ٥,٥

٧ - من البيانات الآتية تنبأ بدرجة أحمد فى المتغير ص إذا كانت درجته فى قياس التوتر (س) ٩٤:

مجـ س = ٢٤٠ مجـ ص = ١٩٥

مجـ س٢ = ٧٥٠ مجـ س ص = ٥١٠

ن = ١٠

٨- يوضح الجدول الآتى العلاقة بين السعادة الزوجية (س) والمستوى التعليمى (ص) لعينة من الأسر .. والمطلوب حساب معامل التوافق بين هاتين الظاهرتين :

ص	أسر سعيدة	أسر متوسطة السعادة	أسر غير سعيدة	مجـ
تعليم عالى	٣٥	١٢٠	٣	٥٠
تعليم متوسط	١٨	١٦	٦	٤٠
أمى	٢	٧	٢١	٣٠
مجـ	٥٥	٣٥	٣٠	١٢٠

٩- الجدول التالى ، يبين عمر أحد النباتات بالأسابيع وطوله بالسنتيمتر

العمر	١	٢	٣	٤	٥	٦	٧
الطول	٥	١٣	١٦	٢٣	٣٣	٣٨	٤٠

والمطلوب إيجاد:

(أ) خط انحدار الطول على العمر.

(ب) الطول عند عمر مقداره ٩ أسابيع.

١٠ - لدراسة العلاقة بين الدخل (س) والعمر بالسنوات (ص) بين عمال أحد المصانع. أخذت عينة مكونة من ٢٠ عامل فأعطت النتائج الآتية:

مجـ س = ١٠   مجـ $س^٢$ = ١٤

مجـ س ص = ٢٩,٢   مجـ ص = ٢٠

مجـ $ص^٢$ = ٨٤

والمطلوب إيجاد:

(أ) خط انحدار الدخل على العمر.

(ب) دخل العامل الذى يبلغ من العمر ٣٠ سنة.

١١- لدراسة العلاقة بين العمر (ص) بالسنوات ومدة الحياة الزوجية (س) بالسنوات كانت لدينا النتائج الآتية:

مجـ س = ٥٠   مجـ ص = ٤٠

مجـ س ص = ٥٨٠   مجـ $س^٢$ = ٥٠٠

مجـ $ص^٢$ = ٨٠٠

والمطلوب حساب:

(أ) معامل الارتباط بين العمر ومدة الحياة الزوجية.

(ب) خط انحدار العمر على مدة الحياة الزوجية.

(جـ) تقدير العمر عندما تصل مدة الحياة الزوجية ٢٠ سنة.

١٢- لدراسة العلاقة بين الدخل (ص) بمئات الدولارات، الاستهلاك (س) بمئات الدولارات فى مدينة نيويورك، أخذت عينة من ٤٠ أسرة فأعطت النتائج الآتية:

مجـ س = ١٠٠ مجـ ص = ١٢٠

مجـ س ص = ٥١٦ مجـ $س^٢$ = ٤١٠

مجـ $ص^٢$ = ٧٢٠

والمطلوب حساب :

(أ) معامل الارتباط بين الدخل والاستهلاك.

(ب) خط انحدار الدخل على الاستهلاك.

(جـ) قيمة الدخل عندما يبلغ الاستهلاك ٧٠٠ دولار.

١٣- لدراسة العلاقة بين الكمية المطلوبة بين الكمية المطلوبة (ص) والسعر (س) بعشرات الدولارات كانت لدينا النتائج الآتية:

مجـ س = ٦٠ مجـ ص = ٧٠

مجـ س ص = ٣٧٤ مجـ $س^٢$ = ٤٠٦

مجـ $ص^٢$ = ٥٣٦ ن = ١٠

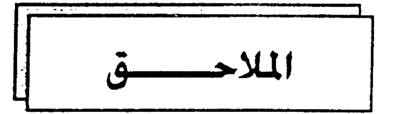
والمطلوب حساب:

(أ) معامل الارتباط بين الكمية المطلوبة والسعر.

(ب) خط انحدار الكمية المطلوبة على السعر.

(جـ) الكمية المطلوبة عندما يصل السعر إلى ٧٠ دولار.

* * * * *


الملاحـــق

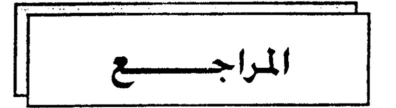
جدول رقم (١)

دلالـة معامـلات الارتبـاط

درجة الحرية	الدلالة	
ن-٢	عند ٠,٠٥	عند ٠,٠١
١	٠,٩٩٧	١,٠٠٠
٢	٠,٩٥٠	٠,٩٩٠
٣	٠,٨٧٨	٠,٩٥٩
٤	٠,٨١١	٠,٩١٧
٥	٠,٧٥٤	٠,٨٧٤
٦	٠,٧٠٧	٠,٨٣٤
٧	٠,٦٦٦	٠,٧٩١
٨	٠,٦٣٢	٠,٧٦٥
٩	٠,٦٠٢	٠,٧٣٥
١٠	٠,٥٧٦	٠,٨٠٨
١١	٠,٥٥٣	٠,٦٧٤
١٢	٠,٥٣٢	٠,٦١١
١٣	٠,٥١٤	٠,٦٤١
١٤	٠,٤٩٧	٠,٦٢٣
١٥	٠,٤٨٢	٠,٦٠٦
١٦	٠,٤٦٨	٠,٥٩٠
١٧	٠,٤٥٦	٠,٥٦٥
١٨	٠,٤٤٤	٠,٥٦١
١٩	٠,٤٣٣	٠,٥٤٩
٢٠	٠,٤٢٣	٠,٥٣٧
٢١	٠,٤١٣	٠,٥٢٦
٢٢	٠,٤٠٤	٠,٥١٥

تابع جدول رقم (١)

درجة الحرية		الدلالة
ن-٢	عند ٠,٠٥	عند ٠,٠١
٢٣	٠,٣٩٦	٠,٥٠٥
٢٤	٠,٣٨٨	٠,٤٩٦
٢٥	٠,٣٨١	٠,٤٨٧
٢٦	٠,٣٧٤	٠,٤٧٨
٢٧	٠,٣٦٧	٠,٤٧٠
٢٨	٠,٣٦١	٠,٤٦٣
٢٩	٠,٣٥٥	٠,٤٥٦
٣٠	٠,٣٤٩	٠,٤٤٩
٣٥	٠,٠٢٥	٠,٣٩٣
٤٠	٠,٣٠٤	٠,٣٩٣
٤٥	٠,٢٨٨	٠,٣٧٢
٥٠	٠,٢٧٣	٠,٣٥٤
٦٠	٠,٢٥٠	٠,٣٢٥
٧٠	٠,٢٣٣	٠,٣٠٢
٨٠	٠,٢١٧	٠,٢٨٣
٩٠	٠,٢٠٥	٠,٢٦٧
١٠٠	٠,١٩٥	٠,٢٥٤
١٢٥	٠,١٧٤	٠,٢٢٨
١٥٠	٠,١٥٩	٠,٢٠٨
٢٠٠	٠,١٣٨	٠,١٨١
٣٥٠	٠,١١٣	٠,١٤٨
٤٠٠	٠,٠٩٨	٠,١٢٨
٥٠٠	٠,٠٨٨	٠,١١٥
١٠٠٠	٠,٠٦٢	٠,٠٨١


المراجـــــع



أولاً - المراجع العربية :

١- أحمد عبادة سرحان: مقدمة فى الإحصاء الاجتماعى، الجزء الأول - جامعة الإسكندرية، كلية التجارة.

٢ - اعتماد علام ويسرى رسلان ، أساسيات الإحصاء الاجتماعى ، دار قطرى بن الفجاءة للنشر والتوزيع ، الدوحة ، قطر ، ١٩٩١ .

٣ - السيد محمد خيرى : الإحصاء فى العلوم النفسية والتربوية والاجتماعية، الطبعة الثالثة، القاهرة، مطبعة دار التأليف، ١٩٦٤ .

٤ - زكريا الشربينى ، الإحصاء اللابارامترى ، مكتبة الانجلو المصرية ، القاهرة ١٩٩٠ .

٥ - عبد اللطيف عبد الفتاح وأحمد محمد عمر: مقدمة الطرق الإحصائية، الطبعة الرابعة ، القاهرة، شركة الطوبجى للطباعة والنشر، ١٩٨١، ١٩٨٢ .

٦ - عبدالرحمن بن محمد سليمان ابو عمه ، أنور أحمد محمد عبدالله ومحمود هندى ، الاحصاء التطبيقى ١٩٩٠ .

٧ - عبد الله عبد الحليم أبو بكر، داود سليمان المدنى وإسماعيل سليمان العوامرى: أساليب البحث الإحصائى، القاهرة ، التجارة والتعاون للطبع والنشر، ١٩٨٤ .

٨ - فاروق عبد العظيم وآخرون : مبادىء الإحصاء الوصفى والتحليل، الإسكندرية ، دار المطبوعات الجامعية، ١٩٨٤ .

٩ - فاروق عبد العظيم أحمد، وبدر الدين المصرى : الإحصاء ، القاهرة، دار الكتب الجامعية.

١٠ - فتحى عبد العزيز أبو راضى : مبادىء الإحصاء الاجتماعى، الجزء الأول، الإسكندرية، دار المعرفة الجامعية.

١١ - محمد سمير إبراهيم وأبو بكر أحمد حسين : أساسيات علم الإحصاء، الجزء الأول، الطبعة الثانية، القاهرة، مكتبة عين شمس، ١٩٦٧ .

١٢ – محمود السيد أبو النيل : الإحصاء النفسى والاجتماعى ومعايير إظهار الشخصية الإسقاطى الجمعى، الطبعة الثانية، القاهرة الجهاز المركزى للكتب الجامعية والمدرسية والوسائل التعليمية ١٩٧٨ .

١٠– مصطفى رزق : الكمبيوتر للمبتدئين ، الطبعة الثانية، أسيوط مكتبة الطليعة، ١٩٨٦ .

١١– يحيى هندام ومحمد الشبراوى على : أساسيات الإحصاء فى البحوث الاجتماعية الطبية، القاهرة، مكتبة النهضة الحديثة .

ثانياً - المراجع الأجنبية:

1– Anderson. T.W. and Stanley L - Sclove, Statistical Analysis of Data. Boston: Houghton Miffin Company. 1978.

2- Blalock. Hubert M, Social Statistics, 2 nd. edition. New York: Mc Graw- Hill Book Company. 1972.

3- Bogue, Donald J. Principles of Demograpny New York: John wiley and Sons. Inc. 1969.

4- Ferman S . Gerald and Jack Levin, Social Science Research: A Handbook for Students. New York: John wiley and Sons, 1975.

5- Felice, L., Statistics : a Tool for Social Research, Blemont, California : Wadswarth Publishing

6- Graham, Alan, Statistics, London : Hodder Headline Plc., 1994

7- Hinkle, Dennis, wiersm, william and Jurs, G .Stephen, Applied Statistics for the Behavioral Sciences. Chicago: Rand Mc Nally College publishing Company. 1979.

8- Kurtz, Norman R. Introduction to Social Statistics. Tokyo: Mc Graw - Hill Book Company. 1983.

9- Lutz, Gene M..Understanding Social Statistics. New York: Macmillan publishing Co.. Inc. 1983.

10- Nie, H. Normam et al.. Statistical Package. For The Social Sciences.New York: Mc Graw - Hill Book Company. 1978.

11- Ott, Lyman. An Introduction to Statistical Methods and Data Analysis - North Scituate. Massachusstss: Duxbury press. 1977.

12- Parsons, Robert. Statisical Analysis: A Decision - Making Approach. New York: Harper and Raw Publihers, 1979.

13- Shryock. Henry S. and Jacob S. Siegel, The Methods and

Materials of Demography.. U. S. Department of Commerce. Vol. 1, 1980.

14- Startup : Richard and Ewyn T. Vhittaker. Introducing Social Statistics. London: George Allen and Unwin, 1982.